UNIVERSIDAD TECNOLÓGICA DE PEREIRA

DOCTORAL THESIS

Probabilistic modeling of tensorial data for enhancing spatial resolution in magnetic resonance imaging

Author:
Hernán Darío Vargas Cardona

Supervisor:
Ph.D Álvaro A. Orozco

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Engineering Doctorate Program
Department of Electrical Engineering

May 9, 2018

Declaration of Authorship

I, Hernán Darío Vargas Cardona, declare that this thesis titled: *Probabilistic modeling of tensorial data for enhancing spatial resolution in magnetic resonance imaging* and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

UNIVERSIDAD TECNOLOGICA DE PEREIRA

Abstract

Faculty of Engineering
Department of Electrical Engineering

Doctor of Philosophy

Probabilistic modeling of tensorial data for enhancing spatial resolution in magnetic resonance imaging

by Hernán Darío Vargas Cardona

In biological tissues such as nerve fiber bundles and muscles, the spontaneous heat motion of water molecules is restricted by obstacles in the fibrous microstructure. Medical imaging uses the principles of Magnetic Resonance Imaging (MRI) to non-invasively measure properties of this motion. MRI provides unique insights about tissue connectivity, which make this procedure one of the key technologies in an ongoing large-scale scientific effort to map the human brain connectome. MRI and diffusion-MRI (dMRI) have been used in several applications: medical procedures, signal processing, computer vision, and neurosciences. Although, current protocols allow fast acquisitions in a different number of slices at several planes, the spatial resolution is not high enough in many cases for clinical diagnosis. The main problem occurs due to hardware limitations in acquisition scanners. Therefore, MRI and dMRI have a difficult compromise between a good spatial resolution and signal noise ratio (SNR). This leads to data acquisitions with low spatial resolution. It becomes in a serious issue for clinical analysis for two main reasons. First, a low spatial resolution in visual data reduces quality in important medical processes such as: diagnosis of diseases, segmentation, anatomical atlas construction, etc. Second, to obtain high resolution images is required a long time acquisition. However, current clinical protocols do not allow a long exposition of radiation (MRI and dMRI) in human subjects. According to the previously established, we propose a probabilistic framework based on stochastic processes (STs) for enhancing spatial resolution in different modalities of magnetic resonance imaging. The proposed methodology is robust to high presence of noise. Also, It ensure physically realizable data and does not modify information necessary for segmentation or clinical analysis in comparison to the methods proposed in literature. The method for resolution enhancement of MRI supports the accurate tissue segmentation of gray matter, white matter and cerebrospinal fluid in anatomical images. For diffusion images (dMRI), the proposed approaches ensure positive definite tensors, one of them generalizes to any order and they avoid the loss of relevant clinical information obtained from Diffusion Tensor Imaging (DTI) and Higher Order Tensors (HOT) such as fractional and generalized anisotropy maps and fiber tracts.

Acknowledgements

I thank to my advisor Álvaro A. Orozco and my co-advisor Mauricio A. Álvarez for their dedication and permanent support.

I thank to *Colciencias*, specifically to the program: *formación de alto nivel para la ciencia, la tecnología y la innovaciíon - Convocatoria 617 de 2013*, for financing my doctorate degree.

Also, I thank to the project: *Desarrollo de un sistema de soporte clínico basado en el procesamiento estocástico para mejorar la resolución espacial de la resonancia magnética estructural y de difusión con aplicación al procediento de ablación de tumores*, funded by *Colciencias* with code 1110-744-55860.

I thank to the members of the automatics research group for their collaboration.

Finally, I thank to my entire family for trusting me.

Contents

List of Figures

List of Tables

To my Grandma...

Chapter 1

Introduction

1.1 Problem Statement

In biological tissues such as nerve fiber bundles and muscles, the spontaneous heat motion of water molecules is restricted by obstacles in the fibrous microstructure. Medical imaging uses the principles of Magnetic Resonance Imaging (MRI) to non-invasively measure properties of this motion [1]. When it is applied to the human brain, this provides unique insights about tissue connectivity, which makes MRI one of the key technologies in an ongoing large-scale scientific effort to map the human brain connectome [2]. Consequently, it is a timely and important topic of research to create mathematical models that infer biologically meaningful parameters from such data. MRI and diffusion-MRI (dMRI) have been used in applications ranging from signal processing [3], computer vision [4], and neurosciences [5]. Although, current clinical protocols allow fast acquisitions in a different number of slices at several planes, the spatial resolution is not high enough in many cases for clinical diagnosis [6]. The main problem occurs due to hardware limitations in acquisition scanners [7]. Therefore, MRI and dMRI have a difficult compromise between a good spatial resolution and signal noise ratio (SNR). This leads to data acquisitions with low spatial resolution. It becomes in a serious issue for clinical analysis for two main reasons. First, a low spatial resolution in visual data reduces quality in important medical processes such as: diagnosis of diseases, segmentation (tissue, nerves and bone), anatomical atlas construction, detailed fiber reconstruction (tractography), brain conductivity models, etc. Second, to obtain high resolution images is required a long time acquisition. However, current clinical protocols do not allow a long exposition of radiation (MRI and dMRI) in human subjects.

In anatomical MRI studies, there are two main types of image acquisitions: MRI-T1 and MRI-T2. T1 imaging is based on the exponential recovery of longitudinal magnetization. T2 imaging is based on exponential loss of signal resulting from purely random spin-spin interactions in the transverse of the XY plane [8]. But, the spatial resolution becomes in a considerable difficulty. Although 3D Fourier acquisition is commonly the suitable procedure for MRI, where a 3D high-resolution (HR) is required, this option is not available in practice for all desired image contrast mechanisms [9]. Also, T1 and T2 are highly affected by several classes of noise: artifacts, physiological noise, subject motion, etc. For this reason, classical methods for processing digital images are not suitable in this context. Spatial resolution of anatomical MRI has been widely studied in the last 20 years [10, 11, 12, 13, 14]. However, these methodologies did not solve the spatial resolution problem satisfactorily. For example, there are practical shortcomings: low performance in presence of high level of noise, sensitivity to parameter modifications, low mathematical robustness, among others [11, 12].

On the other hand, from dMRI it is possible to estimate the apparent diffusivity coefficient (ADC) of water particles within tissue by solving the Stejskal-Tanner formulation [15]. The first attempt to represent this physical phenomena was the Gaussian model proposed by [16], where symmetric and positive definite tensors of rank-2 quantify the direction and orientation of diffusion. From this quantification, it is possible to compute relevant physiological information (i.e. Fractional anisotropy and mean diffusivity). Related to dMRI, the limitation in spatial resolution is bigger than T1 and T2. Current protocols allow a data acquisition with a resolution in a range from 1 and 2 mm^3 for each voxel. Here, the problem is that brain tissue fibers are in micro-metric scale, so, fiber reconstruction models (tractography) can be inaccurate. Usually, tensorial visualization is discrete, where ellipsoids are employed for graphic representation [17]. A considerable number of methods for dMRI interpolation of rank-2 tensors have been proposed in the literature [18, 19, 20, 21, 22, 23]. They have different drawbacks: some methods do not ensure positive definite tensors and they are highly affected by the intrinsic Rician noise added during acquisition [20, 21, 23].

While, rank 2 tensors (Diffusion Tensor Imaging-DTI) is the method of choice to represent dMRI data. DTI fails to represent accurately some complex tissue structures such as: white matter fiber bundles, crossing fibers and bifurcated fibers [24, 25]. To address these limitations in dMRI, several researchers [24, 26, 27, 28] proposed higher order tensor (HOT) models in order to describe diffusion inside complex tissue structures. Higher order tensors are considered as a generalization of matrices to multi-way arrays. Tensorial modeling is an active and challenging topic in applied mathematics, since fundamental concepts from linear algebra, such as the singular value decomposition, do not have unique generalization to higher order, and most generalizations are hard to compute [29]. Again, spatial resolution arises as the main problem to be solved. Also, methods for tensorial interpolation are only valid for rank-2 tensors and a generalization for HOT interpolation is missing [30].

According to the above, it is clear that there are methodological problems related to spatial resolution in magnetic resonance imaging (MRI and dMRI) have not been solved entirely. As was pointed out before, current methods employed in the literature for enhancing resolution in MRI (T1, T2) [10, 11, 12, 13, 14] and dMRI (DTI, HOT) [18, 19, 20, 21, 22, 23, 31] fail in high presence of noise. Also, they do not ensure physically realizable data or do not preserve relevant clinical information, among others. As opposed, probabilistic methods give advantages compared to the state of the art. For example, interpretation of model predictions, robustness to noise data and provide well-founded framework for learning and model selection [32]. In addition it is known that segmentation procedures (structures or type of tissue) are not accurate enough in low resolution images. This is explained because edges, contours, details and strong changes of contrast are not well defined [11, 12]. Therefore, some problems related to clinical applications remain unsolved, due to the low spatial resolution achieved in all MRI modalities. For this reason, there arises the following research question: ¿Is it possible to develop a probabilistic modeling of tensorial data for enhancing spatial resolution in magnetic resonance imaging (MRI-T1, MRI-T2, dMRI-DTI and dMRI-HOT), allowing the improvement of tissue segmentation and preserving morphological properties compared to the same procedure performed in low resolution images?

1.2 Justification

1.2.1 Relevance

An important aspect is that improving the spatial resolution in MRI (anatomical or diffusion) is always done at the expense of either the signal-to-noise ratio (SNR) or the acquisition time. For the last 20 years, super-resolution (SR) algorithms have successfully been applied to MRI data (T1-T2 and dMRI) to increase the spatial resolution of scans, thus facilitating the clinical diagnosis [12, 20, 31]. The variety of application and techniques has grown ever since, especially in the dMRI modality, showing the interest of the community to such post acquisition processing [10, 13, 14, 21, 22, 23, 31]. Also, increasing of spatial resolution allows a better segmentation procedure due to the improvement of key features in MRI such as: high contrast with a filtering stage, edge enhancement, defined contours and detailed transitions [12].

For anatomical MRI (T1 and T2) the super-resolution methodologies are the best approaches to enhance spatial resolution through post-processing of 2D multi-slice images [33]. Currently, clinical protocols allow fast acquisitions of a considerable number of slices in different planes. Although 3D Fourier acquisition is commonly the suitable procedure for MRI, where a 3D high-resolution (HR) is required, this option is not available in practice for all desired image contrast mechanisms [9]. SR approaches use MRI studies with low-resolution (LR) images for enhancing edge definition in each slice [33], and for improving clinical procedures such as tissue segmentation, registration, target detection, tracking and disease diagnosis.

For dMRI (rank-2 and HOT fields), the tensorial interpolation is a feasible methodology to enhance spatial resolution, and takes relevance in the reconstruction of tissue fiber bundles (this procedure is known as tractography). Also, in image registration algorithms or any application where it is required to estimate data among nearby tensors of the field [34]. Diffusion tensors (DTs) have been used in applications ranging from psychometrics [35] and chemometrics [36] to signal processing [3], computer vision [4], and neuroscience [5]. They also provide adequate models for a number of quantities that occur in the context of diffusion imaging. Many practitioners view HOT as a generalization of matrices to multi-way arrays. However, tensors can also be studied in an invariant, coordinate-free notation. Tensor decompositions are an active and challenging topic in applied mathematics, since fundamental concepts from linear algebra, such as the singular value decomposition, do not have unique generalization to higher order, and most generalizations are hard to compute.

1.2.2 Viability

Previous research works [10, 11, 12, 13, 14, 18, 19, 20, 21, 22, 23], related to the enhancement of resolution in MRI anatomical (T1 and T2) and dMRI (Rank-2 and HOT) have proved that it is possible to achieve a better visualization and information extraction, when the spatial resolution is improved. Consequently, better outcomes in clinical procedures are achieved where MRI and dMRI are necessary. For example: pathologies diagnosis, surgical planning of neuro-diseases, atlas construction, among others. Moreover, stochastic processes applied in physical phenomena modeling have demonstrated flexibility, accuracy and robustness. In the medical imaging field, the probabilistic modeling is adequate for enhancing spatial resolution. Nowadays, the probabilistic modeling through stochastic processes

(STs) have become in a recognized and validated machine learning methodology. A ST defines a distribution over functions, and inference taking place directly in the space of functions. The ST is a practical and probabilistic approach to perform supervised learning. It gives advantages compared to the state of the art related to the interpretation of model predictions, robustness to noise data and provides well-founded framework for learning and model selection [32]. Research communities in statistical image processing showed that 2D or 3D high resolution images can be seen as a Gaussian process regression (GPR) problem in MRI (T1 and T2). Generalized Wishart processes (GWP) or multi-output GPs are suitable for rank-2 tensors employed in lower order dMRI representation. Finally, tensorial decomposition processes describe the dMRI-HOT data. In this context, the prediction of the target voxel depends of a learning process with training data. Specifically, the training samples correspond to the low resolution studies (T1, T2, DTI and HOT). Therefore, this framework is able to perform accurate tissue segmentation, and preserves the relevant information in tensorial data, no matter the intrinsic noise added in acquisition of these images.

Also, the research group in automatics (RGA) has successfully developed several research works including medical imaging processing. Specifically, the MRI has been employed in planning of surgical procedures performed in the institute of epilepsy of the eje cafetero in Pereira, Colombia. Likewise, The RGA has developed brain conductivity models through dMRI applied in software for clinical support during deep brain stimulation realized in Parkinson's disease patients. It is worth to noting that this work has the advice and support of the *laboratorio de procesamiento de imagen médica*, affiliated to the Universidad Rey Juan Carlos from Madrid, España.

1.3 List of contributions

We summarize the contributions of this research work in the following items:

1. We introduced a robust probabilistic framework based on stochastic processes (STs) for enhancing spatial resolution in different modalities of magnetic resonance imaging. See sections 2.4, 2.5 and 2.6.

2. We developed a probabilistic methodology for enhancing spatial resolution of anatomical MRI (T1 and T2) based on Gaussian processes. We obtained better segmentation metrics for tissue types (white matter, gray matter, and cerebrospinal fluid) in comparison with low resolution images. See sections 2.4 and 4.1. The mentioned methodology was published in [37].

3. We contributed with a methodology for interpolation of diffusion tensor (DT) fields obtained from dMRI with stationary and non-stationary Multioutput Gaussian Processes (GP). Specifically, we decompose each tensor in six features: three eigen-values and three Euler angles. Next, we interpolate the features as related outputs of the GP. See subsections 2.5.2 and the following publications: [38, 39].

4. We presented a novel methodology for interpolation of DT fields. Instead of the commonly employed Riemannian framework, we propose a stochastic modeling of DT field. We assume that a DT field follows a generalized Wishart process (GWP). In this context, we use it to model the entire DT field $D(x, y, z)$. Then, through approximate Bayesian inference (i.e. Elliptical slice sampling and Markov Chain Monte Carlo methods), we estimate the optimal

parameters of the model. Results of GWP improve to the comparison methods [18, 19] in different validation protocols. See subsections 2.5.1, 4.2.1 and this publication: [40].

5. We applied the diffusion tensors in simulation models for deep brain stimulation. We showed an objective comparison among three kind of volumes which represent the geometry, layers and structure of the head, likewise, the influence of electrical properties of anisotropic tissue and ground positioning. Another worth contribution of this work is the analysis and evaluation of the anisotropy level in relevant brain tissue (i.e. Thalamus-Thal, Subthalamic Nucleus-STN and Substantia Nigra reticulata-SNr) in Colombian patients. For this purpose, we used DT fields obtained from five patients with Parkinson's Disease located in the west-central region of Colombia. See these publications: [41, 42].

6. We proposed a novel methodology to perform interpolation in HOT fields of any order. In this regard, we employ tensor representations and modulate their parameters with GPs, aiming to estimate new data with robustness, considering that GPs are functions of a multi-dimensional input variable. Specifically, we introduce two probabilistic models, that we refer to as the Tucker decomposition process (TDP) and the canonical decomposition process (CDP). Our models are based on the Tucker and canonical decomposition of tensors, respectively. We test the TDP and CDP in 2nd, 4th and 6th rank HOT fields. For rank-2 tensors, we compare against direct interpolation [18], log-Euclidean approach [19], and Generalized Wishart processes [40]. For rank-4 and rank-6 tensors we compare against direct interpolation and raw dMRI interpolation with b-splines [31]. TDP and CDP interpolate accurately the HOT fields, and generalize to any rank. Importantly, the proposed framework safeguards the mandatory constraint of positive definite tensors, and preserve morphological properties such as fractional anisotropy (FA), generalized anisotropy (GA), and tractography. See sections 2.6 and 4.3.

7. We generated two research projects (approved and financed) from this thesis:

   - *Desarrollo de un sistema de soporte clínico basado en el procesamiento estocástico para mejorar la resolución espacial de la resonancia magnética estructural y de difusión con aplicación al procedimiento de ablación de tumores*, financed by Colciencias with code 111074455860. Years: 2017-2018. Total amount: $350.000.000.00 COP.
   - *Spatial resolution enhancement for medical imaging segmentation using stochastic modeling*, financed by Universidad Tecnológica de Pereira with code 6-16-1. Years: 2016-2017. Total amount: $25.000.000.00 COP.

1.4 Aims

1.4.1 General aim

To develop an interpolation framework based on probabilistic modeling of tensorial data for enhancing spatial resolution in magnetic resonance imaging (MRI-T1, MRI-T2, dMRI-DTI, dMRI-HOT). The proposed framework must improve the tissue segmentation and preserving morphological properties compared to the same procedure performed in low resolution images.

1.4.2 Specific aims

1. To develop a stochastic methodology for super-resolution of anatomical MRI (T1 y T2) through scalar stochastic processes for supporting tissue segmentation.

2. To build a stochastic interpolation approach applied in rank-2 tensor fields for enhancing spatial resolution in diffusion tensor images (DTI) and preserving fractional anisotropy of diffusion images. Application to fiber tracts reconstruction.

3. To formulate a probabilistic approach for interpolation of higher order tensors (HOT) fields allowing the enhancement of spatial resolution for construction of generalized anisotropy maps.

Chapter 2

Referential Framework

2.1 Related Work

2.1.1 Anatomical MRI: T1 and T2

Currently, it is possible to achieve fast acquisitions of a considerable number of slices in different planes. However, spatial resolution is not high enough in many cases for clinical diagnosis [6]. Although 3D Fourier acquisition is commonly the suitable procedure for MRI, where a 3D high-resolution is required, this option is not available in practice for all desired image contrast mechanisms [9]. For this reason, the scientific community have been developing super-resolution methodologies to enhance spatial resolution through post-processing of 2D multi-slice images. SR approaches use MRI studies with LR images for enhancing edge definition in each slice [33], and for improving clinical procedures such as tissue segmentation, registration, target detection, tracking and disease diagnosis.

The MRI resolution problem has been widely studied. There are many algorithms and methods that deal with the issue of LR images due to the available hardware and acquisition methods [7]. The earliest attempts to up-sample resolution in medical images were presented in [43, 44, 10], and a method based on inter-slice reconstruction using SR was proposed in [11]. But their performance is reduced in presence of high levels of noise. Also, approaches based on Wiener filter regularization [45], and non-local means filtering [13] have been developed to perform SR images, obtaining improved studies. Current methods for SR images are based in patch learning and regularization. In [46, 47], the authors introduced a dynamic patch scheme applied to MR image reconstruction. The problem with this method resides in the dependence with the number of nearest-neighbors of the patch. In addition, they assume that input LR-MR images are free of noise, despite the intrinsic noise produced by scanners. In [12, 14], sparse methodologies with overcomplete dictionaries were introduced, obtaining high quality MR images. However, there are some drawbacks in [12, 14] due to the complex optimization algorithms. Finally, an automated method based on random forest algorithm was presented in [48], showing that supervised learning can offer convenient results in SR medical images. Clearly, there are unsolved issues related to the segmentation of structures or tissue in anatomical MRI due to low resolution.

2.1.2 diffusion MRI: Rank-2 and higher order tensors (HOT)

Diffusion magnetic resonance imaging has become in the last years in a very relevant topic inside medical imaging and clinical applications [49]. dMRI allows to describe non-invasively the diffusion of water particles in biological tissues using

tensors. Mathematically, a DT associated to a *voxel* is represented with a symmetric and positive definite 3×3 matrix [50]. A 3D grid composed by interconnected diffusion tensors is known as a DT field. These fields are widely employed in multiple clinical applications such as: brain connectivity [50], segmentation of different types of tissue (white matter, gray matter, Cerebrospinal fluid), diagnosis of neurodegenerative diseases, electrical conductivity models for the brain [51], atlas construction, image registration, among others.

Nevertheless, due to the current acquisition protocols and some limitations in MRI machines, it is not possible to acquire dMRI data with enough spatial resolution, for obtaining accurate clinical analysis derived form diffusion images. Usually, dMRI acquired with these protocols have *voxels* whose resolution are in a range from 1 to 2 mm^3. This problem has two main consequences. First, most of applications of DTI fields require higher-resolution data. Otherwise, the accuracy on some medical procedures is not good enough [52]. Second, the fiber bundles of tissue are much smaller than the typical voxel size. Therefore, 3D mapping of brain connectivity (Tractography) may be discontinuous [53]. For these reasons, enhancing of spatial resolution in dMRI through digital processing methods is key for performing precise clinical procedures.

To enhance dMRI data resolution in a tensorial space, interpolation provides an interesting and feasible methodological solution. Rank-2 tensor fields belong to a Riemannian space, where the Riemannian metric is defined by the inner product assigned to each point of this space. With this metric, one can to compute geodesic distances between diffusion tensors and to calculate different statistics in this space [21]. An important condition is keeping the smooth transition of anisotropic features inherent in the given tensor fields (i.e. Fractional anisotropy-FA maps), especially around degenerate points, where at least two of three eigenvalues are equivalent [54].

Some recent works have proposed interpolation methods for rank-2 tensor fields. They developed a variety of mathematical approaches, such as: direct smooth approximation [18] and euclidean approaches, but they do not retain the principal properties of a DTI, i.e positive definite tensors. For this reason, the scientific community has been looking alternative methods for estimating tensor fields that keep the symmetric positive definite (SPD) constraint inside the grid of tensors. [19] presented a Log-euclidean approximation and [21] developed a Riemannian framework achieving important advances in tensor fields geometry, but they lack in smoothness property in presence or high level of noise. [20] presented a b-spline scheme that interpolates SPD tensors with high accuracy using the Riemannian metric. The authors introduced a tensor product of B-splines that minimizes the Riemannian distance between tensors. Following the Riemannian framework, [22] presented Geodesic-loxodromes that can identify isotropic and anisotropic components of the tensor and interpolates each component separately. Finally, alternative methodologies have been posited: a tensor field reconstruction based on eigenvector and eigenvalue interpolation [17], location of degenerated lines in 2-D planar [54], and a feature-based interpolation [23]. However, those methods do not achieve an adequate representation of a rank-2 tensor field obtained from noisy real data and none of them can be generalized to higher order tensor (HOT) fields. Also, [31] proposed the enhancement of resolution in scalar gray-level of dMRI data with conventional interpolation methods. However, [55] showed that scalar interpolation of dMRI does not avoid the undesirable swelling effect, and it is recommended the enhancement of spatial resolution of dMRI in a tensorial space.

Regarding HOT field interpolation, the authors of [56, 34] developed a method

based on tensor subdivision and minimization of two properties (curl and divergence) of the field for interpolation of 4th-order tensors. However, [56, 34] only reported outcomes for rank-4 tensors fields, and the method is not clear an extension to any order. Also, the authors of [29, 30] proposed an approach to perform probabilistic tractography in HOT data. This methodology is based on Finsler geometry. They developed the geometric generalization appropriate for multi-fiber analysis and demonstrated that a HOT field belongs to a Finsler manifold. Although, the Finsler geometry model is able to perform tractography in tensors fields of any order, it has not been established as a interpolation method.

2.2 Basic principles of magnetic resonance imaging (MRI)

2.2.1 Nuclear Magnetic Moment

In clinical MRI the image is formed by the signals from protons in water and lipid. At the atomic level, since a proton is a charged particle which spins around an internal axis of rotation with a given value of angular momentum P, it also has a magnetic moment μ, and therefore can be thought of as a very small magnet with a north and south pole, as shown in Fig. 2.1. The magnetic field generated by the spin of the particle are collinear with the direction of the spin axes and normally it is termed magnetic moment. The strength of the magnetic moment determines the sensitivity of detection in magnetic resonance and it is dependent on the type of nucleus.

Most frequently, the hydrogen nucleus with one proton is the nucleus of choice in MRI because it possesses the strongest magnetic moment and its abundance in organic tissues. The variations in spin angular momentum result from interactions with an applied static magnetic field and electromagnetic radiation. These particles have mass and thus generate angular momentum as they rotate. Positively and negatively charged particles can be regarded as spheres of distributed positive or negative charges, while neutral electrical particles such as the neutron can be thought of as a combination of distributed positive and negative charges. Relevant to MRI, the magnitude of the angular momentum of the proton is quantized and has a single, fixed value. The magnitude of the proton's magnetic moment is proportional to the magnitude of the angular momentum [57]:

$$|\vec{\mu}| = \gamma|\vec{P}|$$

where γ is a constant called the gyromagnetic ratio, and has a value of 267.54 MHz/Tesla for protons. As a result, the magnitude of the magnetic moment has a single fixed value. In absence of an external magnetic field, as shown in Fig. 2.1, the magnitude of magnetic moment of every proton in our bodies is fixed, but the orientation is completely random. Therefore, the net magnetization, i.e. the sum of all the individual magnetic moments in our bodies, is zero.

The situation changes with the application of an external magnetic field $\mathbf{B_0}$. From quantum mechanics, the component of the magnetic moment in the direction of $\mathbf{B_0}$ can have only two possible discrete values, which results in the magnetic moments being aligned at an angle of $54.7°$ with respect to the direction of $\boldsymbol{B_0}$, aligned either in the same direction. The former configuration is termed as the parallel, and the latter as the anti-parallel configuration: note however that the terms parallel and anti-parallel only refer to the $z-component$ of μ, and that μ is actually aligned at an angle with respect to $\mathbf{B_0}$. The relative number of protons in the parallel and anti-parallel configurations depends upon the value of $\mathbf{B_0}$. Protons in the parallel configuration are preferred because it guarantees the lowest energy state. The energy difference ΔE between the two states is shown in Fig. 2.2 and given by [58]:

$$\Delta E = \frac{\gamma h \mathbf{B_0}}{2\pi} \tag{2.1}$$

where h is the Planck's constant $(6.63 \times 10^{-34} J \cdot s)$. To calculate the relative number of protons in each of the two configurations the Boltzmann equation can be used:

$$\frac{N_{anti-parallel}}{N_{parallel}} = e^{-\frac{\Delta E}{kT}} = e^{-\frac{\gamma h \mathbf{B_0}}{2\pi kT}} \tag{2.2}$$

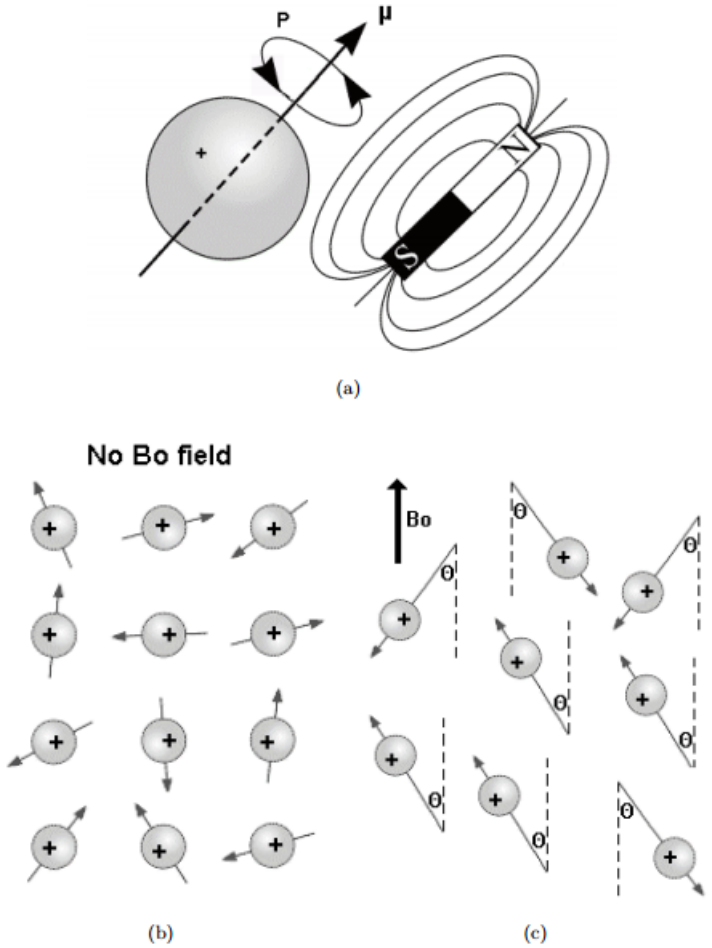


FIGURE 2.1: (a) The internal rotation of a proton creates a magnetic moment, and so the proton acts as a magnet with north and south pole. (b) In the absence of a strong magnetic field, the orientation of the magnetic moments are completely random. (c) When there is a strong magnetic field present the magnetic moment must be align at the angle $\theta = \pm 57°$ with respect to the direction of $\mathbf{B_0}$.

where k is Boltzmann's constant with value $1.38 \times 10^{-23} J/K$, and T is the temperature measured in Kelvin. Since the value of the exponent is very small, a first order approximation $e^{-x} = 1 - x$ can be made:

$$\frac{N_{anti-parallel}}{N_{parallel}} = 1 - \frac{\gamma h \mathbf{B_0}}{2\pi k T} \tag{2.3}$$

The MRI signal depends upon the difference in populations between the two energy levels:

$$N_{parallel} - N_{anti-parallel} = N_{total} \frac{\gamma h \mathbf{B_0}}{2\pi k T} \tag{2.4}$$

where N_{total} is the total number of protons. It is important to note that MRI can detect only the difference $N_{parallel} - N_{anti-parallel}$ and not the total number of protons. As shows equations (2.2) and (2.3), thermal energy causes the energy difference between the two orientations to be minimal, with the two orientations almost equally populated resulting in a net bulk magnetization M. Naturally, the protons

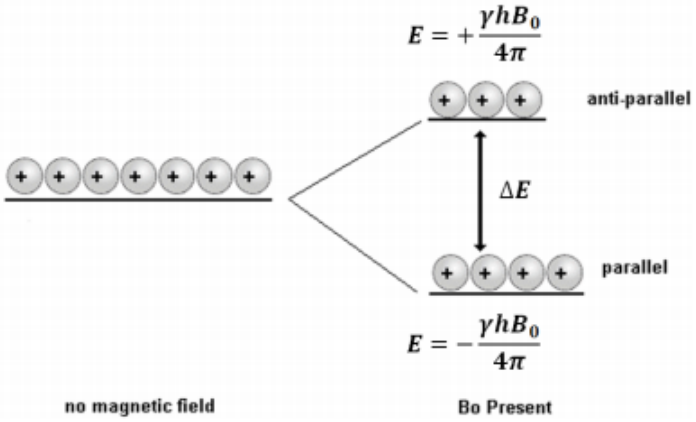


FIGURE 2.2: Proton configurations. (left) in the absence of a strong magnetic field, the energies of all the random orientations of the magnetic moments are the same. (right) When a strong magnetic field is applied, the single energy level splits into two levels, on corresponding to the magnetic moments being in the parallel state, and the other the anti-parallel state. The energy difference between the two states depends upon the value of $\mathbf{B_0}$.

can change from one orientation to another by absorbing or emitting photons with energy equal to the energy difference, as shown in Fig. 2.2.

Classical Precession

Having determined that the proton magnetic moments are all aligned at an angle of $54.7°$ with respect to the direction of $\mathbf{B_0}$. The motion of these magnetic moments can most easily be described using classical mechanics. The $\mathbf{B_0}$ field attempts to aling the proton magnetic moment with itself, and this action create a torque C, given by the cross product of the two magnetic fields:

$$\vec{C} = \vec{\mu} \times \vec{B_0} = i_N |\vec{\mu}\vec{B_0}| sin(\theta)$$

where i_N is a unit vector normal to both $\vec{\mu}$ and $\vec{B_0}$ The direction of the torque $\vec{C}$ shown in Fig.3.3, is tangential to the direction of $\vec{\mu}$ and so causes the proton to "precess" around the axis of the magnetic field, while keeping a constant angle of $54.7°$ between $\vec{\mu}$ and $\vec{B_o}$. To calculate how fast a proton precesses, we use the fact that the torque is defined as the rate of change of the proton's angular momentum:

$$\vec{C} = \frac{d\vec{P}}{dt} = \vec{u} \times \vec{B_0}$$

From Fig. 2.3, the magnitude of the component of the angular momentum which precesses in the plane perpendicular to $\mathbf{B_0}$ is given by $|P|sin(\theta)$. In a short time dt, $\vec{\mu}$ precesses through an angle $d\varphi$ resulting in a change $d\vec{P}$ in the angular momentum. Simple trigonometry gives the relationship that:

$$sin(d\varphi) = \frac{d\vec{P}}{|P|sin(\theta)} = \frac{\vec{C}dt}{|P|sin(\theta)}$$

if $d\varphi$ is small, then we can make the approximation that $sin(d\varphi) = d\varphi$. The angular precession frequency ω, is given by $d\varphi/dt$ and so has a value:

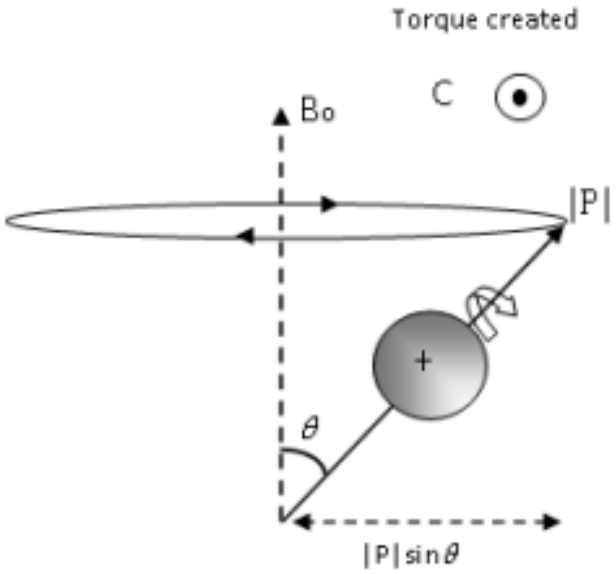


FIGURE 2.3: A proton in a magnetic field. Using classical mechanics, the torque C acting on the magnetic moment, spinning about an internal axis, causes it to precess about the vertical axis **B0**.

$$\omega = \frac{d\varphi}{dt} = \frac{\vec{C}dt}{|\vec{P}|sin(\theta)} = \frac{\vec{u} \times \vec{B_0}}{|\vec{P}|sin(\theta)}$$
$$= \frac{\gamma \vec{P} \times \vec{B_0}}{|\vec{P}|sin(\theta)} = \frac{\gamma |\vec{P}\vec{B_0}|sin(\theta)}{|\vec{P}|sin(\theta)} = \gamma B_0 \tag{2.5}$$

The effect of placing a proton in a magnetic field, therefore, is to cause it to precess around **B0** at a frequency directly proportional to the strength of the magnetic field. This frequency ω_0, is called the Larmor frequency after renowned Irish physicist Joseph Larmor. For hydrogen protons, γ is given as $4257\ Hz/Gauss$. Thus, in a field strength of $7\ T$, the hydrogen proton will precess with a frequency of 297.99 MHz ≈ 300 MHz [59].

Total Magnetization

By superimposing several proton magnetic moments we can represent the net magnetization in a simple vector form. Figure 2.4 shows on the left a representation of several proton magnetic moments, each aligned at $54.7°$ to $\mathbf{B_0}$, each precessing at a frequency ω_0, with slightly more protons in the parallel than anti-parallel state. The total magnetization can be calculated by a simple vector sum of the individual components, and is shown on the right of Fig. 2.4. It can be seen that the net magnetization has only a z-component since the vector sum of the components has only a z-component, since the vector sum of the components on the x- and y-axes are zero. The net magnetization of the sample is defined as $\mathbf{M_0}$ [60]:

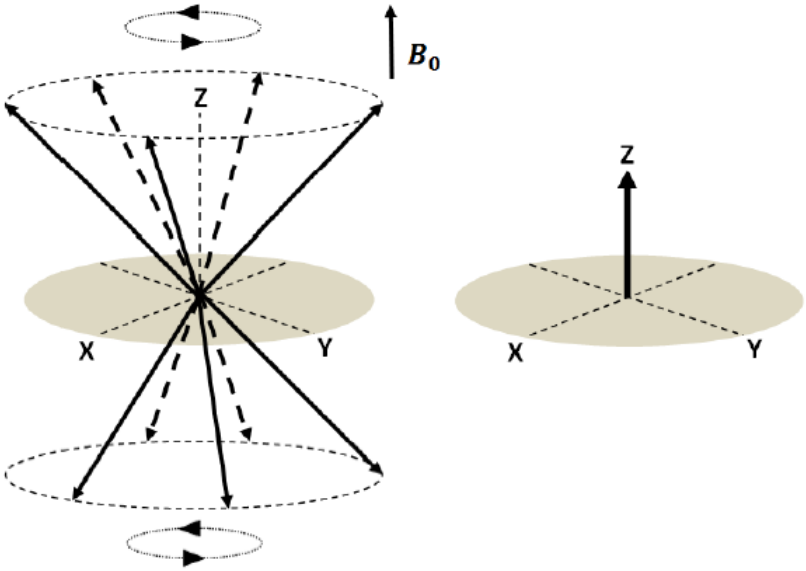


FIGURE 2.4: Magnetization represented by vectors. (left) individual magnetization vectors are randomly distributed around a cone which subtends an angle of $54.7°$ with respect to the $\mathbf{B_0}$ (z) axis. The vector sum of all the individual magnetization vectors (right) is simply a static component in the direction of $\mathbf{B_0}$.

$$\mathbf{M_0} = \sum_{n=1}^{N_{total}} \mu_{z,n} = \frac{\gamma h}{4\pi}(N_{parallel} - N_{anti-parallel}) = \frac{\gamma^2 h^2 \mathbf{B_0} N_{total}}{16\pi^2 kT} \tag{2.6}$$

2.2.2 Effects of Radio Frequency Pulses on Magnetization

The detection of an NMR signal is facilitated by the establishment of a resonance condition. The resonance condition represents a state of alternating absorption and dissipation of energy. Energy absorption is achieved through the application of RF pulses, while energy dissipation is caused by relaxation processes. The energy levels for in a magnetic field, shown in Fig. 2.1, are analogous to energy levels in semiconductors. As with all such a multi-level systems, to obtain an MR signal, energy must be supplied with a specific value ΔE, given by Equation 2.1, to stimulate transitions between the energy levels. The energy is supplied as an electromagnetic (EM), usually as a rediofrequency (RF) field, the frequency (f) of which can be calculated from the Broglie's relationship $\Delta E = hf$:

$$hf = \Delta E = \frac{\gamma h B_0}{2\pi} \Rightarrow f = \frac{\gamma B_0}{2\pi} \; or \; \omega = \gamma B_0 \tag{2.7}$$

By comparing equations (2.7) and (2.5) it can be seen that the frequency of the RF

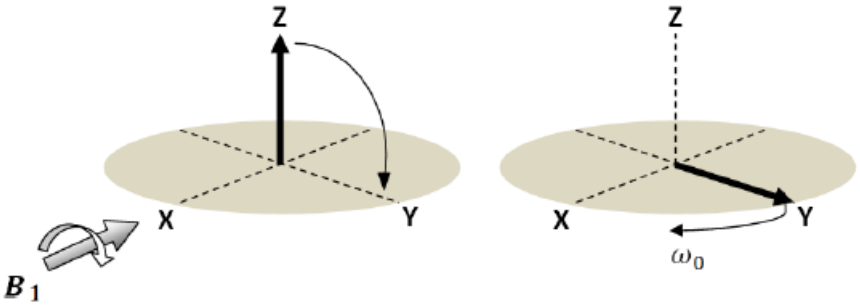


FIGURE 2.5: On the left, application of an RF pulse about the x-axis rotated the magnetization from z-direction towards the y-axis. If the RF pulse strength and duration are chosen to produce 90° pulse, then the magnetization lies directly along the y-axis. When the RF pulse is switched off (right), the magnetization precesses around the z-axis at the Larmor frequency ω_0.

field is identical to the precession frequency. Consider the application of RF radiation at Larmor frequency to a bulk sample of non-magnetic material in an applied static magnetic field. In MRI, the energy is applied as a short RF pulse and it is composed by a coupled electric and magnetic field components [59]. The magnetic field component is denoted by $\mathbf{B_1}$, and it resides in a plane perpendicular to $\mathbf{B_0}$, as shown in Fig. 2.5. Applying the same classical analysis as for proton precession, $\mathbf{B_1}$ produces a torque which causes the net magnetization to rotate towards the xy plane as shown in Fig. 2.5. The consequence of the application of $\mathbf{B_1}$ is to rotate $\mathbf{M}$ by a certain angle away from the $\mathbf{B_0}$ axis. This angle is called the flip angle (α) is defined as the angle through which the net magnetization is rotated [61]. This angle is proportional to both the streght of the applied RF field (measured in Tesla) and the time τ_{B_1}, for which it is applied:

$$\alpha = \gamma B_1 \tau_{B_1}$$

Hence, if $\mathbf{B_1}$ persists for the appropriate duration of time, $\mathbf{M}$ can be made to rotate onto the transverse plane. While in the transverse plane and rotating at the Larmor frequency, $\mathbf{M}$ will induce an NMR signal in the RF receiver coil which is oriented in the transverse plane as shown in Fig. 2.6. This signal can be used to observe the characteristics of $\mathbf{B1}$ in the transverse plane and constitute the basis of MR signal detection, this process constitutes the basis of MR signal detection. The RF pulse that brings $\mathbf{M}$ into the transverse plane is usually referred to as the 90° pulse. A flip angle of 90° results in the maximum value of the $\mathbf{M_y}$ component of magnetization, whereas one of 180° produces no $\mathbf{M_y}$ magnetization but rotates the net magnetization $\mathbf{M_0}$ form +z to the -z axis. The 90° flip angle is very important because the strongest NMR signal is obtained when $\mathbf{M}$ rotates in the transverse plane. The 180° flip angle is primary important in spin-echo imaging techniques where it is used to reverse the direction of $\mathbf{M}$ once it is on the transverse plane [62].

2.2.3 The Basis of MR signal detection

In the most simple case, the MR detector consists of a pair of conductive loops (of copper wire for example) placed close to the patient at an angle of 90° with respect to each other. Faraday's law of induction states that the voltage (V) is induced in each

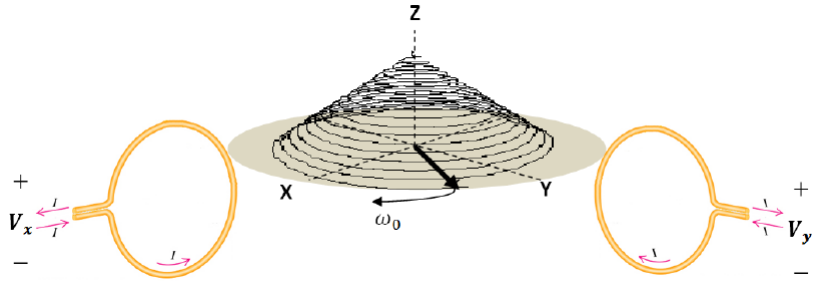


FIGURE 2.6: The MR signal is measured via Faraday induction. Either one or two RF coils can be used, with a voltage being induced across the ends of the conductor loops by the precessing magnetization.

of these loops with a value proportional to the time rate of change of the magnetic flux $d\varphi$:

$$V \propto \frac{d\varphi}{dt}$$

Figure 2.6 shows the situation a short time after $90°$ pulse has been applied about the x-axis: in this case the respective voltages induced in the two coils are given by [63]:

$$\begin{aligned} V_x &\propto M_0\omega_0 sin(\omega_0 t) \\ V_y &\propto -M_0\omega_0 cos(\omega_0 t) \end{aligned} \tag{2.8}$$

It is important to note that the requirement for a time-varying magnetic flux to induce an MR signal is the reason why only magnetization precessing in the xy-plane gives rise to an MR signal. Any z-component of magnetization does not precess and therefore does not induce any voltage.

MR signal intensity

The intensity of the received MR signal is determined by three different factors. First the signal is proportional to the number of protons in the object, from Equation. (2.4). In terms of MRI, as will be seen later, this corresponds to the number of protons in each voxel of the image. The other two factors depend upon the value of the B0 field. From Equation (2.6), the value of M_0 is proportional to B_0. Therefore, a 3 Tesla MRI system has twice the M_0 of a 1.5 Tesla system. Aditionally, from Equation (2.8), the induced voltage is proportional to the precession frequency, which in turn is proportional to B_0. Overall, therefore the MR signal is proportional to the square of the B_0 field, one of the reason of why there is a such a strong drive towards higher field MRI systems.

2.2.4 Relaxation

In a presence of a strong magnetic field B_0, the equilibrion magnetization state corresponds to a z-component, M_z, equal to M_0 and transverse components, M_x and M_y, equal to zero. Application of an RF pulse creates a non-equilibrium state by adding energy to the system. After the pulse has been switched off, the system must relax back to termal equilibrium. The phenomenon of MRI relaxation is similar to the application of an impulse voltage pulse to an RC electrical circuit, where the

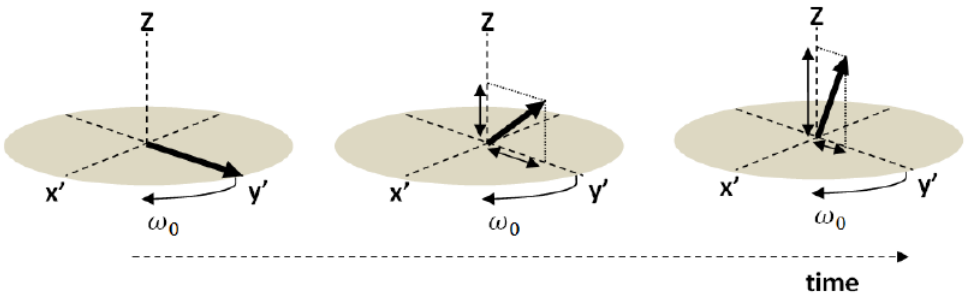


FIGURE 2.7: (left) magnetization vector after a $90°$ RF pulse about the x-axis. (centre) T1 and T2 relaxation of the magnetization a certain time after the pulse has been applied results in an increased Mz component and reduced M_y component, respectivley. (right) After further time, the M_z and M_y components have almost returned to their equilibrium values of M_0 and zero, respectively.

circuit produces time-varying voltages across the lumped elements, the values of which return in time to their values prior the pulse being applied, this process being characterized by certain time-constants. After the application of a $90°$ RF pulse, **M** rotates in the transverse plane at the Larmor frequency and gradually decays to zero as shown in Fig. 2.7.

There are two relaxation times which govern the return to equilibrium of the z-component, and the x- and y- components, respectively. These are referred to as T1-relaxation (which affects only z-magnetization) and T2-relaxation (which affects only x-and y-magnetization). These are also called spin-lattice T1 and spin-spin T2 relaxation [8]. MR relaxation is described mathematically by first order differential equations known as the Bloch equations. Solutions of these equations yields the relation of the M_z component at a time t with the flip angle of an RF pulse after being applied, given by:

$$M_z(t) = M_0 cos(\alpha) + (M_0 - M_0 cos(\alpha))(1 - e^{-\frac{t}{T_1}})$$

For example, after a $90°$ pulse the value of M_z is given by:

$$M_z(t) = M_0(1 - e^{-\frac{t}{T_1}})$$

Different tissues have different values of T_1, and deseased tissues often have substantially altered T_1 relaxation time compared to healthy tissue, and these differences form the basis for introducing contrast into the MR image.

The second relaxation time, T_2, governs the return of the M_x and M_y components of magnetization to their thermal equilibrium values of zero. If an RF pulse of arbitrary flip angle is applied along the x-axis, the value of M_y at time t after the RF pulse is given by:

$$M_y(t) = M_0 sin(\alpha) e^{-\frac{t}{T_2}}$$

As is the case for T1 relaxation times, different tissues in the body have different values of T2, and these can also be used to differentiate between healthy and diseased tissues in clinical images [8].

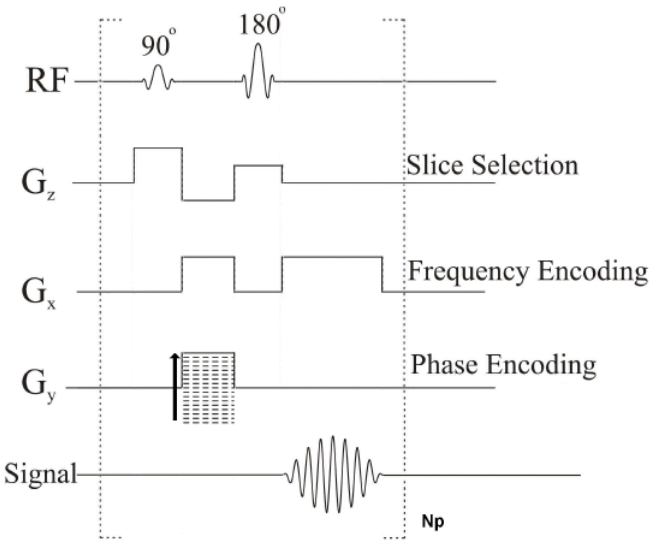


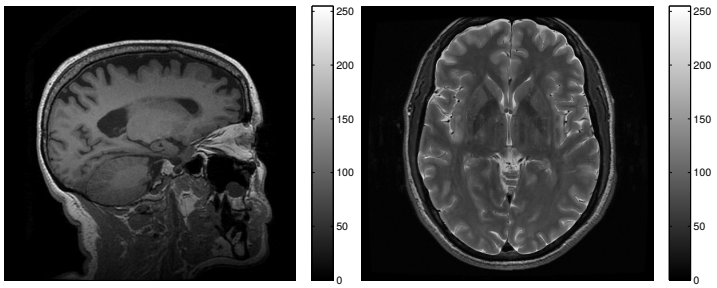
FIGURE 2.8: Pulse sequence diagrams for imaging sequences. An RF pulse is applied, various gradients are turned on and off. Individual steps in image formation can be considered independently in terms of slice selection(RF and G_z), phase-encoding (G_y) and frequency encoding (G_x).

2.2.5 Image Acquisition

The process of image formation can be broken down into three separate, independent components, slice selection, phase-encoding and frequency-encoding. An overall imaging 'pulse sequence' is shown in Fig. 2.8. The transmitter line indicates when an RF pulse is applied, and the length and power of the pulse are adjusted to give an indicated flip angle For each gradient line, the height of the gradient pulse indicated its strength, and the polarity(positive or negative) indicates which direction current is owing through the particular gradient coil. The entire sequence of RF pulse and three gradients has to be repeated a number of times (N_p typically is between 128 and 512) to build-up a two dimensional data-set, with the arrow next to the phase encoding gradient indicating that different values are used for each repetition of the sequence [61].

2.2.6 Summary

- Spinning protons, frequently referred to as just spins, can be represented by a small magnetization vector.
- A slight excess number of protons align with the main external magnetic field $\mathbf{B_0}$. The number in excess is proportional to the magnitude of $\mathbf{B_0}$. The total magnetic field of the excess protons is the equilibrium magnetization $\mathbf{M_0}$.
- Spinning protons wobble or precess about the external field $\mathbf{B_0}$. The resonance frequency is called the Larmor frequency ω_0 and is proportional to $\mathbf{B_0}$.
- Radio frequency pulse: a magnetic field with a transverse component oscillating at the Larmor frequency can be used to rotate the magnetization vector. If a RF pulse with flip angle α is applied to the equilibrium magnetization $\mathbf{M_0}$, then after the RF pulse the new longitudinal and transverse components will be $M_z = M_0 cos(\alpha)$ and $M_{xy} = M_0 sin(\alpha)$, respectively .



(a) (b)

FIGURE 2.9: Examples of anatomical MRI. (a) is a T1 image and (b) is a T2.

- NMR signal: a transverse component M_{xy} will precess about $\mathbf{B_0}$ with the Larmor frequency, thus inducing a measurable signal in a properly tuned coil.
- Spin-Lattice Relaxation: The process whereby energy absorbed by the excited protons or spins is released back into the surrounding lattice, reestablishing thermal equilibrium.
- T1 Relaxation: Spin-Lattice relaxation. The exponential recovery of longitudinal (aligned with $\mathbf{B_0}$) magnetization. M_z returns to $\mathbf{M_0}$.
- Spin-Spin Relaxation: The temporary and random interaction between two excited spins that causes a cumulative loss in phase resulting in an overall loss of signal. Also known as transverse or T2 relaxation.
- T2 Decay: The exponential loss of signal resulting from purely random spin-spin interactions in the transverse or XY plane.
- Magnetic resonance imaging (MRI) is a medical technique used in radiology to obtain anatomical images of healthy and pathological tissues. MRI scanners use strong magnetic fields and radio-waves to form the images of the body using the signal information from protons in water and lipid. In MRI studies, there are two main types of image acquisitions: MRI-T1 and MRI-T2. T1 imaging is based on the exponential recovery of longitudinal magnetization. T2 imaging is based on exponential loss of signal resulting from purely random spin-spin interactions in the transverse of the XY plane [8]. Both T1 and T2 are spin echo studies, and they are needed to adequately evaluate a tissue in MRI protocols. The figure 2.9 shows an example of T1 and T2 images.

2.3 Diffusion Magnetic Resonance Imaging (dMRI)

2.3.1 Theory of diffusion

Diffusional processes in the liquid state are governed by the Bloch-Torrey equation:

$$\frac{\partial M_+}{\partial t} = -i\omega_0 M_+ - i\gamma \mathbf{r} \cdot G M_+ - \frac{M_+}{T_2} + D\nabla^2 M_+, \tag{2.9}$$

where $M_x = M_x + iM_y$ is the complex representation of transverse magnetization, $\mathbf{r}$ is the position vector, G is the linear magnetic field gradient, γ is the gyromagnetic ratio, ω_0 is the Larmor frequency, T_2 is the spin-spin relaxation time, and

D is the apparent diffusion coefficient. For a spin-echo experiment, Eq. (2.9) can be solved by using the :

$$M_+(\mathbf{r}, t, G) = A(t)\exp(-i\omega_0 t - \frac{M_+}{T_2} - i\gamma \mathbf{r}\cdot F) \tag{2.10}$$

where

$$\begin{aligned} F(t) &= \int_0^t G(t')dt' - 2\Theta\left(t - \frac{TE}{2}\right)\int_0^{TE/2} G(t')dt' \\ &\approx \mathbf{g}F(t) \end{aligned} \tag{2.11}$$

and $\Theta(x)$ is the Heaviside step function, which is equal to unity when its argument is positive, and 0 otherwise. In the approximation, we have assumed that the direction of the gradients, $\mathbf{g}$, is constant throughout the experiment, so that all of the time dependence is in their magnitude. Therefore, for experiments in which the diffusion gradients are much larger than others, $\mathbf{g}$ can be taken as the direction of the diffusion sensitizing gradients, which can be expressed in terms of the polar angle θ and azimuthal angle ϕ as

$$\mathbf{g} = \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} = \begin{bmatrix} sin\theta\ cos\phi \\ sin\theta\ sin\phi \\ cos\theta \end{bmatrix} \tag{2.12}$$

Inserting Eq. (2.10) into Eq. (2.9), the resulting first-order differential equation yields the Stejskal-Tanner formula for diffusive attenuation:

$$S = S_0\exp(-\gamma^2\delta^2G^2(\Delta - \delta/3)D) = S_0\exp(-bD(g)) \tag{2.13}$$

2.3.2 Rank-2 tensors representation

Diffusion Magnetic Resonance Imaging (dMRI) studies the diffusion of water particles in the human brain. The first attempt for modeling diffusion can be described with a rank-2 tensor given by a symmetric positive definite 3×3 matrix proportional to the covariance of a Gaussian distribution [15].

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix}$$

For water, the diffusion tensor (DT) is symmetric, so that $D_{ij} = D_{ji}$, where $i, j = x, y, z$. The diffusion tensor for each voxel of the dMRI is calculated using the Stejskal-Tanner formulation [15]:

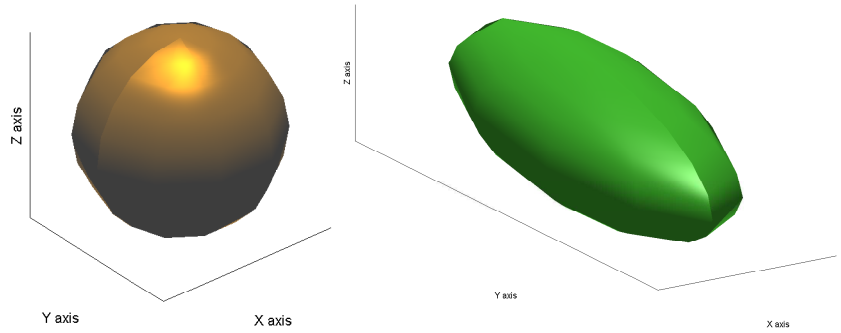
$$S_k(\mathbf{z}) = S_0(\mathbf{z})e^{-b\hat{\mathbf{g}}_k^\top D(\mathbf{z})\hat{\mathbf{g}}_k}, \tag{2.14}$$

where S_k is the k^{th} dMRI, S_0 is the reference image, $\mathbf{z}$ corresponds to spatial coordinates $[x, y, z]^\top$, $\hat{\mathbf{g}}_k$ is the gradient vector, and b is the diffusion coefficient. At least 7 dMRI measurements are necessary for each slice ($k = 0, 1, ..., 7$). Usually, DT fields are estimated from (2.14) using least squares.

Rank-2 DTs have been visualized by constructing the ellipsoid given by:

$$\mathbf{r}^\top D^{-1}\mathbf{r} = C \tag{2.15}$$

where $\mathbf{r}^\top = [x, y, z]$ is the position vector, and C is a constant with the units of time. Therefore, the resulting shape is a level surface of the expression on the left side of (2.15), and it is possible to show by diagonalization that these surfaces are ellipsoids. Figure 2.10 shows an isotropic and anisotropic DT.



(a) Isotropic DT (b) Anisotropic DT

FIGURE 2.10: Isotropic and anisotropic rank-2 diffusion tensors

2.3.3 Higher Order Tensors (HOT)

To provide a formal basis for the modeling of highly structured diffusion, it is proposed an extension to the Bloch-Torrey equation to include a phenomenological diffusion term with a higher-rank (> 2) Cartesian tensor:

$$\frac{\partial M_+}{\partial t} = -i\omega_0 M_+ - i\gamma \mathbf{r} \cdot G M_+ - \frac{M_+}{T_2} + \sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\cdots\sum_{i_l=1}^{3} D_{i_1,i_2\ldots i_l} g_{i_1} g_{i_2} \ldots g_{i_l} \nabla^2 M_+. \tag{2.16}$$

Here $D_{i_1 i_2 \ldots i_l}$ are the components of the Cartesian rank-l tensor, and g_{i_1} through g_{i_l} are components of the gradient vector. Substitution of Eq. (2.10) into equation (2.16) yields a generalized Stejskal-Tanner formula [24]:

$$log(S) = log(S_0) - b \sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\cdots\sum_{i_l=1}^{3} D_{i_1,i_2\ldots i_l} g_{i_1} g_{i_2} \ldots g_{i_l}. \tag{2.17}$$

This equation makes it possible to calculate all the components of the DT of general rank by means of a simple multilinear regression [28]. If l is an odd number, then this relation implies that

$$D(-g) = -D(g),\ is\ odd. \tag{2.18}$$

However, since negative diffusion coefficients are nonphysical, l is forced to be an even number. Therefore, from Eq. (2.17) we arrive at the condition for antipodal symmetry of the diffusivities:

$$D(-g) = D(g),\ is\ even, \tag{2.19}$$

A general rank-l Cartesian tensor has 3^l terms, which is a very large number for higher ranks. For example, a rank-10 tensor will have 59049 components. However,

symmetries provide a very significant reduction in the number of distinct components. This follows from the realization that $D_{i_1 i_2 \ldots i_l}$ is a totally symmetric tensor. Total symmetry is due to the fact that this tensor links the components of the same vector to a scalar $(D(g))$. For example, in the case of $l = 2$,

$$D(g) = \sum_{i=1}^{3}\sum_{j=1}^{3} D_{ij} g_i g_j = \sum_{i=1}^{3}\sum_{j=1}^{3} D_{ji} g_j g_i, \tag{2.20}$$

implies $D_{ij} = D_{ji}$, since it is true for all vectors $\mathbf{g}$. A similar analysis for the general l case yields

$$D_{i_1 i_2 \ldots i_l} = D_{i_l \ldots i_1 i_2}. \tag{2.21}$$

This symmetry reduces the number of distinct elements to

$$N_l = \begin{bmatrix} l+2 \\ 2 \end{bmatrix} = \frac{(l+1)(l+2)}{2}, \tag{2.22}$$

which is only 66 for $l = 10$ case. To use Eq. 2.17 to derive the distinct components of the rank-l DT, it is needed to know how many times a given element is repeated. This is the multiplicity of each element, and it is denoted with the letter μ. Knowing the multiplicity of every unique element, Eq. (2.17) is replaced by

$$log(S) = log(S_0) - b \sum_{k=1}^{N_l} \mu_k D_k \prod_{p=1}^{l} g_k(p), \tag{2.23}$$

where D_k is the k-th unique element of the tensor, and $g_k(p)$ is the component of the gradient direction specified by the p-th index of the k-th unique element of the generalized DT. The multiplicity of a component of a rank-l tensor is given by:

$$\mu = \frac{l!}{n_x! n_y! n_z!} \tag{2.24}$$

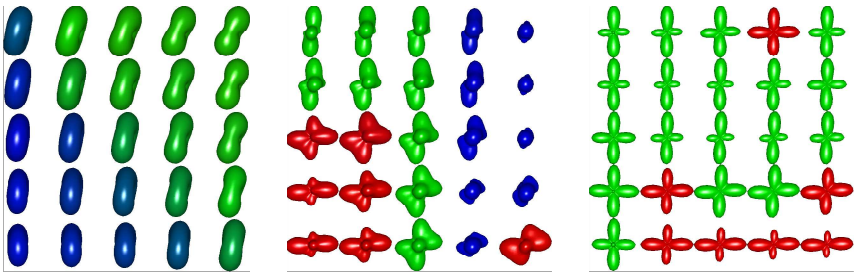
From a whole dMRI study, it is possible to estimate a grid of interconnected and related tensors known as tensor field. A HOT has a discrete graphical representation defined by parametrized surfaces known as glyphs [24]. Figure 2.11 shows examples of HOT fields of rank-2,4 and 6.

2.4 Probabilistic model for enhancing resolution of MRI (T1-T2)

2.4.1 Gaussian Process

A Gaussian Process (GP) is a collection of random variables, any finite number of which have a joint Gaussian distribution [64]. A GP is completely defined by its mean function, $m(\mathbf{x})$, and covariance function, $k(\mathbf{x}, \mathbf{x}')$, such that

$$f(\mathbf{x}) \sim \mathcal{N}\left(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\right),$$



(a) (b) (c)

FIGURE 2.11: Examples of HOT fields: (a) rank-2, (b) rank-4, and (c) rank-6. RGB colors indicate the principal direction of the diffusion tensor: right-left (RED), anterior-posterior (GREEN) and ventral-inferior (BLUE)

or equivalently,

$$f(\mathbf{x}) \sim \mathcal{GP}\left(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\right),$$

where $f(\mathbf{x})$ is the intensity value of pixel $\mathbf{x}$. In supervised learning, the squared exponential kernel is commonly employed as covariance function, and it is given by

$$k(\mathbf{x}, \mathbf{x}') = \sigma^2 \exp\left(-\frac{||\mathbf{x} - \mathbf{x}'||^2}{2\theta^2}\right), \tag{2.25}$$

where θ and σ^2 are the length-scale and the variance hyperparameters, respectively. Be a noisy image with a set of pixels and intensities $\{(\mathbf{x}_i, y_i)\}_{i=1}^n$, where its intensities follow the standard linear regression model $y_i = f_i + \varepsilon$, with $f_i = f(\mathbf{x}_i)$ and Gaussian noise $\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$. The joint distribution of the training intensities, $\mathbf{y}$, and the test intensities, $\mathbf{y}_*$, is given by

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}_* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0}_* \end{bmatrix}, \begin{bmatrix} \mathbf{K}_{\mathbf{y},\mathbf{y}} + \sigma_n^2\mathbf{I} & \mathbf{K}_{\mathbf{y},\mathbf{y}_*} \\ \mathbf{K}_{\mathbf{y},\mathbf{y}_*}^\top & \mathbf{K}_{\mathbf{y}_*,\mathbf{y}_*} + \sigma_n^2\mathbf{I}_* \end{bmatrix}\right),$$

where $\mathbf{K}_{\mathbf{y},\mathbf{y}_*}$, $\mathbf{K}_{\mathbf{y},\mathbf{y}}$ and $\mathbf{K}_{\mathbf{y}_*,\mathbf{y}_*}$ denote the matrix covariances between training and test points, training points, and test points, respectively. Here, the conditional distribution for test points is given by [64]

$$\mathbf{y}_* | \mathbf{X}, \mathbf{y}, \mathbf{X}_* \sim \mathcal{N}\left(\bar{\mathbf{y}}_*, \mathrm{cov}(\mathbf{y}_*)\right),$$

with

$$\begin{aligned} \bar{\mathbf{y}}_* &= \mathbf{K}_{\mathbf{y},\mathbf{y}_*}^\top [\mathbf{K}_{\mathbf{y},\mathbf{y}} + \sigma_n^2\mathbf{I}]^{-1}\mathbf{y}, \\ \mathrm{cov}(\mathbf{y}_*) &= \mathbf{K}_{\mathbf{y}_*,\mathbf{y}_*} + \sigma_n^2\mathbf{I}_* - \mathbf{K}_{\mathbf{y},\mathbf{y}_*}^\top [\mathbf{K}_{\mathbf{y},\mathbf{y}} + \sigma_n^2\mathbf{I}]^{-1}\mathbf{K}_{\mathbf{y},\mathbf{y}_*}. \end{aligned}$$

In Figure 2.12, we show the graphical model for a GP adapted for single images (T1 and T2) using a 3×3. Here, the inputs are the intensities of neighbor pixels (predictors), and the output is the intensity of the target pixel we want to predict.

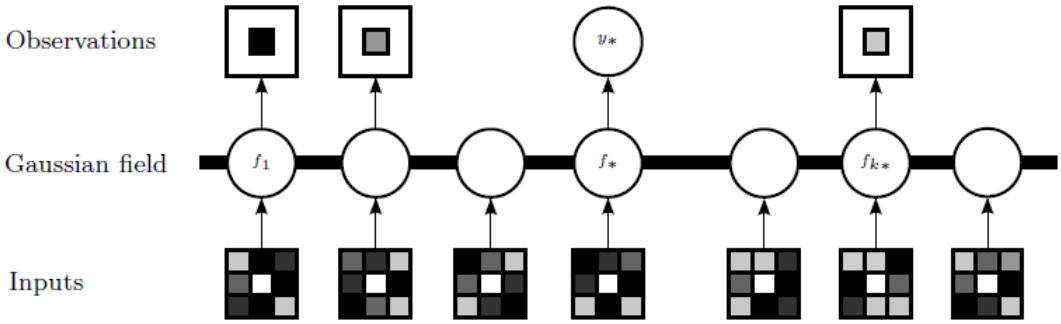


FIGURE 2.12: Graphical model for a GP applied to a single super-resolution image using a 3×3 patch. Squares represent observed pixels and circles represent unknown Gaussian fields. The inputs are the intensities of neighbor pixels (predictors), and the output are the intensity of target pixels we want to predict.

2.5 Probabilistic models for DTI interpolation

2.5.1 Generalized Wishart Process

We begin with the Wishart distribution, which defines a probability density function over a symmetric positive definite matrix. Let S be a $p \times p$ symmetric positive definite matrix of random variables. Let V be a (fixed) positive definite matrix of size $p \times p$. Then, if $\nu \geq p$, S has a Wishart distribution with ν degrees of freedom if it has a probability density function given by:

$$S = \frac{|S|^{\nu-p-1}}{2^{\nu p/2}|V|^{\nu/2}\Gamma_p(\frac{\nu}{2})} e^{-\frac{1}{2}\operatorname{trace}(V^{-1}S)},$$

where $|\cdot|$ is the determinant and Γ_p is the multivariate gamma function:

$$\Gamma_p\left(\frac{\nu}{2}\right) = \pi^{\frac{p(p-1)}{4}} \prod_{j=1}^{p} \Gamma\left(\frac{\nu}{2} + \frac{1-j}{2}\right)$$

Following this notion and according to the definition given in [65], a generalized Wishart process (GWP) is as a collection of symmetric positive definite random matrices indexed by an arbitrary and high dimensional dependent variable $\mathbf{z}$. In DTI fields, the dimension is $p = 3$ because diffusion tensors are represented by 3×3 matrices, and the indexed variable refers to position coordinates $\mathbf{z} = [x, y, z]^\top$. Assume 3ν independent Gaussian process functions $u_{id}(\mathbf{z}) \sim \mathcal{GP}(0, k)$, for $i = 1, ..., \nu$ and $d = 1, 2, 3$, where $k(\mathbf{z}, \mathbf{z}')$ is the covariance or kernel function for the GP. Given a set of input vectors $\{\mathbf{z}\}_{n=1}^{N}$, the vector $(u_{id}(\mathbf{z}_1), u_{id}(\mathbf{z}_2), ..., u_{id}(\mathbf{z}_N))^\top \sim \mathcal{N}(\mathbf{0}, K)$, being K an $N \times N$ Gram matrix with entries $K_{ij} = k(\mathbf{z}_i, \mathbf{z}_j)$. If we define $\hat{\mathbf{u}}_i(\mathbf{z}) = (u_{i1}(\mathbf{z}), u_{i2}(\mathbf{z}), u_{i3}(\mathbf{z}))^\top$ and L as the lower Cholesky decomposition of a $p \times p$ scale matrix V, such that $LL^\top = V$, for each input position $\mathbf{z} = [x, y, z]^\top$, the diffusion tensor $D(\mathbf{z})$ follows a Wishart distribution,

$$D(\mathbf{z}) = \sum_{i=1}^{\nu} L\hat{\mathbf{u}}_i(\mathbf{z})\hat{\mathbf{u}}_i^\top(\mathbf{z})L^\top \sim \mathcal{GWP}_p(\nu, V, k(\cdot, \cdot)), \tag{2.26}$$

In this case, we use the squared exponential kernel (see equation (2.25)).

Bayesian inference for DTI field learning

In order to perform DTI interpolation, we first need to compute the posterior distribution for the variables in the model. For a DTI field, we assume a prior given by a Generalized Wishart process

$$p\left(D(\mathbf{z})\right) \sim \mathcal{GWP}_3(\nu, V, k(\cdot,\cdot)) = \sum_{i=1}^{\nu} L\hat{\mathbf{u}}_i(\mathbf{z})\hat{\mathbf{u}}_i^{\top}(\mathbf{z})L^{\top}. \tag{2.27}$$

For the likelihood function, we assume each element from the diffusion tensor data follows an independent Gaussian distribution with the same variance σ^2. This leads to a likelihood with the following form:

$$p(S|u, L, \nu) \propto \prod_{i=1}^{N} \exp\left(-\frac{1}{2\sigma^2}\|S(\mathbf{z}_i) - D(\mathbf{z}_i)\|_{frob}^2\right),$$

where $S(\mathbf{z})$ is the known initial DTI field with low resolution, $D(\mathbf{z})$ is constructed from equation (2.27), and Frobenius norm is given by

$$\|\mathbf{X}\|_{frob}^2 = \text{trace}\left(\mathbf{X}^T\mathbf{X}\right).$$

The purpose is to infer the posterior probability of $D(\mathbf{z})$ given a known tensorial data set $S(\mathbf{z}) = \{S(\mathbf{z}_1), S(\mathbf{z}_2), ..., S(\mathbf{z}_N)\}$, being N the number of data in the initial DTI field. We first compute the posterior of the relevant variables in equation (2.27) including the vector of all GP function values $\mathbf{u}$, length-scale hyperparameter of the GP kernel function θ, the lower Cholesky decomposition of the scale matrix L, such that $LL^{\top} = V$, and the degrees of freedom ν. Given a GWP prior for the model and the likelihood function, the posterior distributions can be computed by

$$p(\mathbf{u}|\theta, L, S) \propto p(S|\mathbf{u}, L, \nu)p(\mathbf{u}|\theta), \tag{2.28}$$
$$p(\theta|\mathbf{u}, L, S) \propto p(\mathbf{u}|\theta, L, D)p(\theta), \tag{2.29}$$
$$p(L|\mathbf{u}, \theta, S) \propto p(S|\mathbf{u}, L, \nu)p(L). \tag{2.30}$$

We use Markov chain Monte Carlo algorithms to sample in cycles. We employ Metropolis-Hastings to sample θ from (2.29), and the elements of scale matrix L from (2.30). To sample $\mathbf{u}$ from (2.28), we employ elliptical slice sampling [66]. We choose $\nu = 5$ through cross-validation. We set a log-normal prior on θ, a spherical Gaussian prior on elements of L and the prior $p(\mathbf{u}|\theta) \sim \mathcal{N}(\mathbf{0}, K_B)$ is a Gaussian distribution with $3\nu N \times 3\nu N$ block diagonal covariance matrix K_B, formed using 3ν of the K matrices.

DTI field interpolation through GWP modeling

Once we find the posterior distributions over all relevant variables for the model, we can compute the posterior distribution for $D(\mathbf{z}_*)$ in a new spatial position $\mathbf{z}_* = [x_*, y_*, z_*]^{\top}$. First, we have to infer the distribution over all unknown GP function values $\mathbf{u}_*$ in $\mathbf{z}_*$, where $\mathbf{u}_*$ is a vector with elements given by $u_{id}(\mathbf{z}_*)$. The joint distribution over $\mathbf{u}$ and $\mathbf{u}_*$ is given by,

$$\begin{bmatrix}\mathbf{u}\\ \mathbf{u}_*\end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix}K_B & A^{\top}\\ A & I_p\end{bmatrix}\right)$$

If $\mathbf{u}_*$ and $\mathbf{u}$ have p and q elements respectively, A is a $p \times q$ matrix that represents the covariances between $\mathbf{u}_*$ and $\mathbf{u}$ for all pairs of training and validation data, this is $A_{ij} = k_i(\mathbf{z}_*, \mathbf{z}_j)$ for $i + (i-1)N \leq j \leq iN$, and 0 otherwise. I_p is a $p \times p$ identity matrix. Using the properties of a Gaussian distribution, and conditioning on $\mathbf{u}$, we obtain:

$$p(\mathbf{u}_*|\mathbf{u}) \sim \mathcal{N}\left(AK_B^{-1}\mathbf{u}, I_p - AK_B^{-1}A^\top\right) \tag{2.31}$$

From values of $\mathbf{u}_*$ obtained from (2.31), and using equation (2.26), we can construct $D(\mathbf{z}_*)$.

2.5.2 Multi-Output Gaussian Process

Feature-based scheme

First, we decompose each tensor of the field as [23]:

$$\mathbf{D} = \mathbf{E}\Lambda\mathbf{E}^\top, \tag{2.32}$$

where Λ is a diagonal matrix of eigenvalues and the three columns of the matrix $\mathbf{E}$ correspond to the three eigenvectors of the diffusion tensor $\mathbf{D}$,

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \ \mathbf{E} = \begin{bmatrix} \nu_{11} & \nu_{12} & \nu_{13} \\ \nu_{21} & \nu_{22} & \nu_{23} \\ \nu_{31} & \nu_{32} & \nu_{33} \end{bmatrix}$$

We find the Euler angles with the elements of eigenvectors obtained in (2.32),

$$\begin{aligned} \alpha &= \arctan 2(\nu_{12}, \nu_{11}) \\ \beta &= \arctan 2(-\nu_{13}, \sqrt{\nu_{11}^2 + \nu_{12}^2}) = \arctan 2(-\nu_{13}, \sqrt{\nu_{23}^2 + \nu_{33}^2}) \\ \gamma &= \arctan 2(\nu_{23}, \nu_{33}) \end{aligned}$$

Being $\arctan 2(a, b)$ the four-quadrant arctangent of the real arguments a and b. We group the six features from each tensor in a single vector $([\ln \lambda_1, \ln \lambda_2, \ln \lambda_3, \alpha, \beta, \gamma]^\top)$ and we index the field in space coordinates $\mathbf{x} = (x, y, z)$.

Multi-Output GP for DTI interpolation

Multi-output Gaussian processes (MOGP) describes J outputs $\{f_j(\mathbf{x})\}_{j=1}^{J}$, by convolution integrals of latent functions $\{u_q^i(\mathbf{x})\}_{q=1,i=1}^{Q,R_q}$, with kernels $\{G_{j,q}^i(\mathbf{x}-\mathbf{z})\}_{j=1,q=1,i=1}^{J,Q,R_q}$,

$$f_j(\mathbf{x}) = \sum_{q=1}^{Q}\sum_{i=1}^{R_q}\int G_{j,q}^i(\mathbf{x}-\mathbf{z})u_q^i(\mathbf{z})d\mathbf{z}.$$

We assume latent functions $u_q^i(\mathbf{x})$ as independent Gaussian processes with covariance $k_q(\mathbf{x}, \mathbf{x}')$, and $f_j(\mathbf{x})$ form a joint Gaussian process with covariance function $k_{j,j'}(\mathbf{x}, \mathbf{x}')$ where $j, j' = 1, \ldots, J$, given by

$$\sum_{q=1}^{Q}\sum_{i=1}^{R_q}\int\int G_{j,q}^i(\mathbf{x}-\mathbf{z})G_{j',q}^i(\mathbf{x}'-\mathbf{z}')k_q(\mathbf{z}, \mathbf{z}')d\mathbf{z}d\mathbf{z}'.$$

If we assume that $G^i_{j,q}(\mathbf{x}-\mathbf{z}) = a^i_{j,q}\delta(\mathbf{x}-\mathbf{z})$, being $\delta(\mathbf{x})$ the Dirac delta function, we have a particular case for the covariance function known as the linear model of coregionalization (LMC). The covariance $k_{j,j'}(\mathbf{x},\mathbf{x}')$ simplifies to

$$k_{j,j'}(\mathbf{x},\mathbf{x}') = \sum_{q=1}^{Q}\sum_{i=1}^{R_q} a^i_{j,q}a^i_{j',q}k_q(\mathbf{x},\mathbf{x}') = \sum_{q=1}^{Q} b^q_{j,j'}k_q(\mathbf{x},\mathbf{x}'),$$

where $b^q_{j,j'} = \sum_{i=1}^{R_q} a^i_{j,q}a^i_{j',q}$. We use Gaussian kernels for constructing the LMC. Once we develop the LMC, we employ the classical inference method based on maximum likelihood to find the posterior distribution over the outputs $f^J_j(\mathbf{x})$. The reader is referred to [67], for a detailed explanation of MOGP.

In the context of DT fields, the outputs are the features from tensors indexed in $\mathbf{x} = x, y, z$:

$$f^{J=6}_{j=1}(\mathbf{x}) = [\ln\lambda_1(\mathbf{x}), \ln\lambda_2(\mathbf{x}), \ln\lambda_3(\mathbf{x}), \alpha(\mathbf{x}), \beta(\mathbf{x}), \gamma(\mathbf{x})]^\top$$

After we interpolate the features with MOGP, it is necessary to reconstruct the interpolated tensors using:

$$\mathbf{T} = \mathbf{R}\Lambda'\mathbf{R}^\top,\ where:$$

$$\Lambda' = \begin{bmatrix} \exp(\lambda'_1) & 0 & 0 \\ 0 & \exp(\lambda'_2) & 0 \\ 0 & 0 & \exp(\lambda'_3) \end{bmatrix},\ and$$

$$\mathbf{R} = \begin{bmatrix} \cos\alpha\cos\beta & \cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma & \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma \\ \sin\alpha\cos\beta & \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma \\ -\sin\beta & \cos\beta\sin\gamma & \cos\beta\cos\gamma \end{bmatrix}$$

Multi-Output GP with non-stationary kernel

We propose a non-stationary kernel described as the sum of separable kernels, which is expressed as the sum of products of two covariance functions. One computes the dependence between outputs and is independent of the input data x, and other kernel models the dependency between input data, independently of the set functions $\{f_d(\mathrm{x})\}$, given for the function $k_q(\mathrm{x},\mathrm{x}')$.

The authors of [68] proposed a non-stationary kernel for identification and characterization of multidimensional change surfaces. A change surface consists of a convex combination of latent functional regimes,$f_1, ..., f_r$ The transition between any two functions is considered a change surface. On the other hand, we define a non-stationary kernel as an expressive mixture according to the model proposed by [68]: The kernel function has the following form:

$$k(\mathrm{x},\mathrm{x}') = \sum_{i=1}^{r}\sigma(w_i(\mathrm{x}))k_i(\mathrm{x},\mathrm{x}')\sigma(w_i(\mathrm{x}')), \tag{2.33}$$

where $w_i(\mathrm{x}) : \mathbb{R}^P \longrightarrow \mathbb{R}^1$ is the weighting function, with P the dimensional input. The expressiveness of this function determines how many changes can occur in the data. $w_i(\mathrm{x}) = \sum_{j=1}^{v} a_j\cos(\boldsymbol{\omega}_j\mathrm{x} + \mathrm{b}_j)$. $\sigma(z) : \mathbb{R}^1 \longrightarrow [0,1]$, is the warping function, that is computed as a convex combination over the weighting function $\sigma(w_i(\mathrm{x})) = \exp(w_i(\mathrm{x}))/\sum_{i=1}^{r}\exp(w_i(\mathrm{x}))$, $\sum_{i=1}^{r}\sigma(w_i(\mathrm{x})) = 1$ inducing a partial discretization over latent functions. This function produces non-stationarity, since it

depends of the input variable x, and $k_i(\mathrm{x}, \mathrm{x}')$ can be any stationary kernel. It is expected that latent functions have different kernel structures or a same form with different hyperparameters, given by (2.34).

$$w_i(\mathrm{x}) = \sum_{j=1}^{v} a_j \cos(\boldsymbol{\omega}_j \mathrm{x} + \mathrm{b}_j) \tag{2.34}$$

2.6 Probabilistic models for HOT interpolation

2.6.1 Proposed approach for tensor interpolation

A tensor is a geometric or physical object specified by a set of coefficients $\mathcal{T}_{i_1 i_2 \ldots i_l}$ of a multi-linear form $\mathcal{T} = \phi(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_l) \in \mathbb{K}^{I_1 \times I_2 \times \ldots \times I_l}$ of l vector arguments $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_l$ written in some orthonormal basis, where $\mathbb{K}$ may refer to $\mathbb{R}$ (real) or $\mathbb{C}$ (complex). The number l is known as the order or rank of the tensor and each vector argument has an independent (may be different) dimensionality. Alternatively, a tensor can be represented in several forms employing vectorial or matrix approximations:

$$\mathcal{T} \sim \mathbf{M}\left(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_m\right),$$

being $\mathbf{M} \in \mathbb{K}^{I_1 \times I_2 \times \ldots \times I_l}$ any vectorial or matrix decomposition of $\mathcal{T}$, and $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_m$ parameters of the given representation. Following this notion, our main goal is to develop probabilistic models (PM) over tensors indexed by an independent variable $\mathbf{z} = [z_1, z_2, ..., z_J]^\top$, being J the dimensionality of $\mathbf{z}$. For example, if $\mathbf{z}$ refers to spatial coordinates, then, $\mathbf{z} = [x, y, z]^\top$ and $J = 3$. The PM can be seen as probability distributions over a tensor field, this is, a grid of interconnected and related tensors. Furthermore, such probability distributions allow the interpolation of new tensor data for any input locations $(\mathbf{z}_*)$, according to the following definition:

$$\mathcal{T}(\mathbf{z}) \sim \mathcal{M}\left(\boldsymbol{\alpha}_1(\mathbf{z}), \boldsymbol{\alpha}_2(\mathbf{z}), ..., \boldsymbol{\alpha}_m(\mathbf{z})\right), \tag{2.35}$$

where $\mathcal{M}$ is a tensor representation, and $\boldsymbol{\alpha}_1(\mathbf{z}), \boldsymbol{\alpha}_2(\mathbf{z}), ..., \boldsymbol{\alpha}_m(\mathbf{z})$ are free parameters that depend on the mathematical definition of $\mathcal{M}$. The probabilistic nature arises because the parameters $\boldsymbol{\alpha}_1(\mathbf{z}), \boldsymbol{\alpha}_2(\mathbf{z}), ..., \boldsymbol{\alpha}_m(\mathbf{z})$ are realizations of stochastic processes. Specifically to define $\mathcal{M}$, we employ the canonical [69] and the Tucker decomposition [70] of tensors to construct the probabilistic model. The motivation of using both tensorial decompositions is for the simplicity in their representations. It is not necessary to model a complex object such a tensor, but simpler mathematical arrays: scalars, vectors and matrices. This allows to model a tensor field through stochastic methods (i.e. Gaussian processes) that modulate those scalars, vectors and matrices in function of an independent variable $\mathbf{z}$.

2.6.2 Canonical decomposition of a tensor

Any tensor can always be decomposed (possibly non-uniquely) as:

$$\mathcal{T} = \sum_{i=1}^{r} \lambda_i \mathbf{u}_i \otimes \mathbf{v}_i \otimes \cdots \otimes \mathbf{w}_i, \tag{2.36}$$

where $\mathbf{u}_i \in \mathbb{R}^{I_1}$, $\mathbf{v}_i \in \mathbb{R}^{I_2}$, ... $\mathbf{w}_i \in \mathbb{R}^{I_l}$ are unitary vectors, $\lambda_i \in \mathbb{R}^+$ are generalized eigenvalues, and $\otimes$ denotes the outer or Kronecker product. The tensor rank,

rank($\mathcal{T}$), is the smallest integer r such that this decomposition holds exactly. The rank of $\mathcal{T} = [t_{j_1 \dots j_l}] \in \mathbb{K}^{I_1 \times \dots \times I_l}$ is defined as:

$$\text{rank}(\mathcal{T}) := \min \left\{ r \middle| \mathcal{T} = \sum_{i=1}^{r} \lambda_i \mathbf{u}_i \otimes \mathbf{v}_i \otimes \cdots \otimes \mathbf{w}_i \right\},$$

If in (2.36), we have $\mathbf{u}_i = \mathbf{v}_i = \cdots = \mathbf{w}_i$ for every i, then we call it a symmetric outer product decomposition, yielding $I_1 = I_2 = I_l = n$, (being n the dimension of the tensor) and a symmetric rank, $\text{rank}_s(\mathcal{T})$:

$$\text{rank}_s(\mathcal{T}) := \min \left\{ s \middle| \mathcal{T} = \sum_{i=1}^{s} \lambda_i \mathbf{y}_i \otimes \cdots \otimes \mathbf{y}_i, \right\},$$

where $\mathbf{y}_i \in \mathbb{R}^n$ are unitary vectors. Henceforth, we will adopt the following notation:

$$\mathbf{y}^{\otimes l} = \mathbf{y} \otimes \cdots \otimes \mathbf{y}, \ l \text{ copies} \tag{2.37}$$

In addition, to complete symmetry, some applications demand additional constraints, such as positivity definite tensors. Regarding this, a tensor $\mathcal{T}$ of even order is positive definite if and only if the smallest generalized eigenvalue $\lambda_{\min}$ of $\mathcal{T}$ is positive [71]. The definition of symmetric rank is relevant because of the following proposition [72]:

1. *Let $\mathcal{T} \in \mathbb{K}^{n \times \dots \times^l n}$. Then there exist $\mathbf{y}_1, ..., \mathbf{y}_s \in \mathbb{K}^n$, such that*

$$\mathcal{T} = \sum_{i=1}^{s} \lambda_i \mathbf{y}_i^{\otimes l}. \tag{2.38}$$

The above proposition establishes that a symmetric and positive definite tensor can be represented as the superposition of outer (Kronecker) products of s unitary vectors $\mathbf{y}_i \in \mathbb{K}^n$ scaled by the positive generalized eigenvalues $\lambda_i \in \mathbb{R}^+$, $i = 1, ..., s$. The outer product decomposition has often been regarded synonymously as the data analytic models CANDECOMP [69] and PARAFAC [73] where the decomposition is used to analyze multi-way psychometric data.

2.6.3 Canonical decomposition process (CDP)

The CANDECOMP is a superposition of outer products of scaled-vectors. This decomposition represents a symmetric tensor in s positive scalars λ_i and s unitary vectors $\mathbf{y}_i$ whose number of elements depends of the tensor dimensionality. The main advantage of canonical decomposition is due to the easy reconstruction of the tensor, as we can see in equation (2.38). Also, it is worth noting that parameters of this decomposition are simple objects: scalars and vectors, that can be easily described through independent Gaussian processes.

According to the formulation given in equations (2.35) and (2.38), we propose a stochastic approach for tensorial interpolation. Let us define $\boldsymbol{\lambda}(\mathbf{z}) = \{\lambda_1, ..., \lambda_s\}$ as the eigenvalues vector. Following our general approach, the idea would be to index each λ_i by the spatial variable $\mathbf{z}$. Since the values of $\boldsymbol{\lambda}$ should remain positive, we transform the elements λ_i with a $\log$ function. Then, we assume that $\log(\lambda_i)$ follows a Gaussian process. Once we obtain the posterior of $\log \boldsymbol{\lambda}$, we apply the $\exp$ function to recover $\boldsymbol{\lambda}$. Also, for the entries in the unitary vectors $\mathbf{y}_i$ ($i = 1, ..., s$), we

assume each element y_{ji}, $(j = 1, ..., n)$ follows an independent Gaussian process. We normalize each $\mathbf{y}_i$ for ensuring unitary vectors. We refer to this process by the name of the *canonical decomposition process (CDP)*:

$$\mathcal{T}(\mathbf{z}) \sim \mathcal{CDP}(\boldsymbol{\lambda}(\mathbf{z}), \mathbf{y}_i(\mathbf{z}), s, l) = \sum_{i=1}^{s} \lambda_i(\mathbf{z}) \mathbf{y}_i(\mathbf{z})^{\otimes l} \tag{2.39}$$

where, $\log \lambda_i(\mathbf{z}) \sim GP(\boldsymbol{\mu}, k(\mathbf{z}, \mathbf{z}'))$ and $y_{ji}(\mathbf{z}) \sim GP(\mathbf{0}, k(\mathbf{z}, \mathbf{z}'))$. We use an squared exponential kernel (see equation (2.25)), for constructing the covariance of the GPs [32].

2.6.4 Tucker decomposition of a Tensor

Consider $\mathcal{T} \in \mathbb{K}^{I_1 \times I_2 \ldots \times I_l}$ and $\mathbf{A}^{(1)} \in \mathbb{K}^{J_1 \times I_1}$, $\mathbf{A}^{(2)} \in \mathbb{K}^{J_2 \times I_2}$ and $\mathbf{A}^{(l)} \in \mathbb{K}^{J_l \times I_l}$. Then, the Tucker mode-1 product $\mathcal{T} \cdot_1 \mathbf{A}^{(1)}$, mode-2 product $\mathcal{T} \cdot_2 \mathbf{A}^{(2)}$ and mode-l product $\mathcal{T} \cdot_l \mathbf{A}^{(l)}$ are defined by

$$\left(\mathcal{T} \cdot_1 \mathbf{A}^{(1)}\right)_{j_1 i_2 \ldots i_l} = \sum_{i_1=1}^{I_1} \mathcal{T}_{i_1 i_2 \ldots i_l} A^{(1)}_{j_1 i_1}, \ \forall j_1, i_2, ..., i_l,$$

$$\left(\mathcal{T} \cdot_2 \mathbf{A}^{(2)}\right)_{i_1 j_2 \ldots i_l} = \sum_{i_2=1}^{I_2} \mathcal{T}_{i_1 i_2 \ldots i_l} A^{(2)}_{j_2 i_2}, \ \forall i_1, j_2, ..., i_l,$$

$$\left(\mathcal{T} \cdot_l \mathbf{A}^{(l)}\right)_{i_1 i_2 \ldots j_l} = \sum_{i_l=1}^{I_l} \mathcal{T}_{i_1 i_2 \ldots i_l} A^{(l)}_{j_l i_l}, \ \forall i_1, i_2, ..., j_l.$$

A Tucker decomposition of a cubic tensor $(I_1 = I_2 = I_l = n)$ $\mathcal{T} \in \mathbb{K}^{n \times \ldots \times^l n}$ is a decomposition of $\mathcal{T}$ of the form [70]:

$$\mathcal{T} = \mathcal{C} \cdot_1 \mathbf{A}^{(1)} \cdot_2 \cdots \mathbf{A}^{(l-1)} \cdot_l \mathbf{A}^{(l)}, \tag{2.40}$$

in which $\mathcal{C} \in \mathbb{K}^{n \times \ldots \times^l n}$ is known as the core tensor, and $\mathbf{A}^{(1)}$, $\mathbf{A}^{(2)}$,..., $\mathbf{A}^{(l)} \in \mathbb{K}^{n \times R}$ $(R \leq n)$ are matrices with column unitary vectors. If the decomposed tensor is symmetric and positive definite, $\mathbf{A}^{(1)} = \mathbf{A}^{(2)} = ... = \mathbf{A}^{(l)}$. For an $l-$order tensor, equation (2.40) is rewritten as follows:

$$\mathcal{T} = \mathcal{C} \cdot_1 \mathbf{A} \cdot_2 \cdots \mathbf{A} \cdot_l \mathbf{A} = \left[\mathbf{A} \otimes^l \mathbf{A}\right] \operatorname{vec} \mathcal{C}, \tag{2.41}$$

where vec is a operator that transforms a tensor into a vector.

2.6.5 Tucker Decomposition Process (TDP)

The Tucker decomposition is defined by a set of outer products of matrices multiplied by a core tensor. The advantage of Tucker is that it guarantees an exact decomposition, while CANDECOMP may be approximated in some cases. Also, when the tensor is symmetric, the size of representation matrix $\mathbf{A}$ does not depend of the rank, which is a remarkable aspect, because higher orders fields do not increase considerably the time necessary for executing the learning stage.

Based on equations (2.35) and (2.41), we propose another model for tensorial interpolation that we call *Tucker decomposition process* (TDP). Let $\mathcal{T}(\mathbf{z})$ be a random

field of tensors. We say that $\mathcal{T}(\mathbf{z})$ follows a TDP according to:

$$\mathcal{T}(\mathbf{z}) \sim \mathcal{TDP}(\mathcal{C}, \mathbf{A}(\mathbf{z}), l) = \mathcal{C} \cdot_1 \mathbf{A}(\mathbf{z}) \cdot_2 \cdots \mathbf{A}(\mathbf{z}) \cdot_l \mathbf{A}(\mathbf{z}) = \left[\mathbf{A}(\mathbf{z}) \otimes^l \mathbf{A}(\mathbf{z})\right] \operatorname{vec} \mathcal{C}, \quad (2.42)$$

where $\mathcal{C}$ is a $l-$order symmetric core tensor, and $\mathbf{A}$ is a $n \times R$ matrix with column unitary vectors. The variable n is the tensorial dimension, and R is the degree of the decomposition ($R \leq n$). In our probabilistic model, we assume that each element of $\mathbf{A}$ follows an independent GP indexed by $\mathbf{z}$. Again, we normalize each column vector of $\mathbf{A}$ for ensuring unitary vectors. Also, we establish that the unique elements of the core tensor $\mathcal{C}$ are random variables sampled from a spherical multivariate Gaussian distribution. The number of unique elements of a tensor depends on its order l. For example if $l = 4$, we have $E_l = 15$ unique elements in a 4_{th} order tensor. The prior distributions over the parameters in the *TDP* are given by $\mathbf{A}(\mathbf{z})$ with elements $A_{ij}(\mathbf{z}) \sim \mathcal{GP}(\mathbf{0}, k(\mathbf{z}, \mathbf{z}'))$ for $i, j = 1, 2, 3$; $\operatorname{vec} \mathcal{C} \sim \mathcal{N}(\mathbf{0}, c^2\mathbf{I})$, with c^2 the common variance for the elements in $\operatorname{vec} \mathcal{C}$.

2.6.6 Bayesian inference for TDP and CDP

For TDP and CDP, we follow the classical Bayesian approach for finding the posterior parameters:

$$posterior \propto prior \times likelihood$$

Given a finite set of higher order tensors $\mathcal{X}(\mathbf{Z}) = \{\mathcal{D}(\mathbf{z}_i)\}_{i=1}^N$, obtained from solving the Stejskal-Tanner formula for different input locations $\mathbf{z}_i$ ($\mathbf{Z} \in \mathbb{R}^{n \times N}$ is a matrix that contains all spatial locations of the training set, and N is the number of training data), we use Bayesian inference to compute the posterior distribution for the HOT field:

$$p(\mathcal{T}(\mathbf{z})|\mathcal{X}(\mathbf{Z})) \propto p(\mathcal{T}(\mathbf{z}))p(\mathcal{X}(\mathbf{Z})|\mathcal{T}(\mathbf{z})).$$

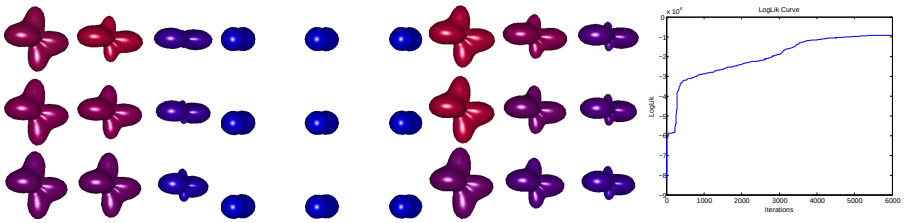
We use the TDP or the CDP as the prior for $p(\mathcal{T}(\mathbf{z}))$, and for the likelihood function, we assume each element from the HOT data follows an independent Gaussian distribution with the same variance σ^2. This leads to a likelihood with the form

$$p(\mathcal{X}(\mathbf{Z})|\mathcal{T}(\mathbf{z})) \propto \prod_{i=1}^{N} \exp\left(-\frac{1}{2\sigma^2}\|\mathcal{X}(\mathbf{z}_i) - \mathcal{T}(\mathbf{z}_i)\|_F^2\right),$$

where $\|\mathcal{A} - \mathcal{B}\|_F$ is the tensorial Frobenius distance of order l given by

$$\|\mathcal{A} - \mathcal{B}\|_F = \left(\sum_{i_1,\ldots,i_l}^{3} \left(\mathcal{A}_{i_1,\ldots,i_l} - \mathcal{B}_{i_1,\ldots,i_l}\right)^2\right)^{1/2}. \quad (2.43)$$

Posterior distributions for the TDP are computed for matrix $\mathbf{A}(\mathbf{z})$, the length-scale parameter θ of the squared exponential kernel (for which a log-normal prior is used), and the core tensor $\mathcal{C}$. We use Markov chain Monte Carlo algorithms to sample in cycles. Metropolis-Hastings [74] is used to sample the posterior of θ, and for the elements of the core tensor $\mathcal{C}$. To sample $\mathbf{A}(\mathbf{z})$, we employ elliptical slice sampling [66]. We set $R = n = 3$ no matter the rank of the HOT field. For the CDP, we employ elliptical slice sampling for obtaining the posterior of $\boldsymbol{\lambda}(\mathbf{z})$ and $\mathbf{y}_i(\mathbf{z})$. We set $s = 8$ when the rank $l = 2$, $s = 10$ when $l = 4$, and $s = 12$ when $l = 6$.



(a) Training data (b) Prior (c) Posterior (d) LogLik curve

FIGURE 2.13: Learning process for a 3×3 rank-4 HOT field. (a) is the training set (low spatial resolution field), (b) is the initial field obtained from the TDP prior, (c) is the posterior field obtained after 6000 iterations, and (d) is the learning curve given by the log-likelihood.

We obtain initial values of relevant parameters sampling from the priors. For MCMC methods, we employ 7000, 9000, 11000 cycles for rank-2, 4 and 6 respectively, taking 1300 for the burn-in stage. The goal of the burn-in is to guarantee statistical independence among samples. For Metropolis-Hastings algorithm, we employ a Gaussian function as proposal distribution with $\sigma^2 = 0.001$. For elliptical slice sampling, we set a learning rate parameter (lr) depending of the rank. For rank-2 tensors we set $lr = 0.001$, for rank-4 and 6 we set $lr = 0.0001$. Coefficients of the tensors are scaled in the range -5 to 5. Figure 2.13 shows an example of the learning process for the TDP.

2.6.7 HOT prediction with the TDP and CDP

Once we learn the posterior distribution for all the parameters, we compute the predictive distribution for $p(\mathcal{T}(\mathbf{z}_*)|\mathcal{T}(\mathbf{z}), \mathbf{z}_*)$, in a new spatial position $\mathbf{z}_* = [x_*, y_*, z_*]^\top$.

First, we have to infer the distribution over all unknown GP function values from $\mathbf{A}(\mathbf{z}_*)$ for the TDP and $\boldsymbol{\lambda}(\mathbf{z}_*)$, $\mathbf{y}_i(\mathbf{z}_*)$ for the CDP. If we vectorized all elements of $\mathbf{A}(\mathbf{z})$, $\mathbf{A}(\mathbf{z}_*)$; $\boldsymbol{\lambda}(\mathbf{z})$, $\boldsymbol{\lambda}(\mathbf{z}_*)$, $\mathbf{y}_i(\mathbf{z})$, $\mathbf{y}_i(\mathbf{z}_*)$, we obtain two vectors $\mathbf{u}$ and $\mathbf{u}_*$ with $p = nRN$ and $q = nRN_v$ (for TDP) or $p = nsN$ and $q = nsN_v$ (for CDP) elements respectively. N is the number of training data and N_v is the number of validation data. The joint distribution over $\mathbf{u}$ and $\mathbf{u}_*$ is given by,

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{u}_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K_B & B^\top \\ B & K^* \end{bmatrix}\right).$$

Here, K_B is a $nRN \times nRN$ (TDP) or $nsN \times nsN$ (CDP) block diagonal covariance matrix, where each block is a $N \times N$ Gram matrix K with entries $K_{ij} = k(\mathbf{z}_i, \mathbf{z}_j)$, being $k(\cdot, \cdot)$ the squared exponential kernel. If $\mathbf{u}_*$ and $\mathbf{u}$ have p and q elements respectively, B is a $p \times q$ matrix that represents the covariances between $\mathbf{u}_*$ and $\mathbf{u}$ for all pairs of training and validation data, this is $B_{ij} = k_i(\mathbf{z}_*, \mathbf{z}_j)$ for $i + (i-1)N \leq j \leq iN$, and 0 otherwise. K^* is a $p \times p$ Gram matrix with entries $K^*_{ij} = k(\mathbf{z}_{*i}, \mathbf{z}_{*j})$. Being $\mathbf{z}_*$, the spatial coordinates of the test data. Using the properties of a Gaussian distribution, and conditioning on $\mathbf{u}$, we obtain:

$$p\left(\mathbf{u}_*|\mathbf{u}\right) \sim \mathcal{N}\left(BK_B^{-1}\mathbf{u}, K^* - BK_B^{-1}B^\top\right) \tag{2.44}$$

From the mean value for $\mathbf{u}_*$ obtained from $p(\mathbf{u}_*|\mathbf{u})$, we organize $\mathbf{A}(\mathbf{z}_*)$ or $\boldsymbol{\lambda}(\mathbf{z}_*)$ and $\mathbf{y}_i(\mathbf{z}_*)$. Then, we compute $\mathcal{T}(\mathbf{z}_*)$ using equations (2.39) and (2.42) for CDP and TDP respectively.

Chapter 3

Materials and Methods

3.1 Materials

3.1.1 MRI dataset

Images of the head were acquired on a General Electric Signa HDxt 3.0T MR scanner using the body coil for excitation, and an 8-channel quadrature brain coil for reception. Subjects were positioned supine with the arms down. The head was positioned in the head-neck coil (MRI). Imaging was performed using an isotropic 3D T1-weighted and 3D T2-weighted Spoiled Gradient Recalled (SPGR) sequence with a repetition time of $10.024\ ms$, an echo time of $4.56\ ms$, an inversion time of $600\ ms$ 1 number of excitations (NEX), an acquisition matrix of 512×512, a resolution of $1 \times 1 \times 1\ mm^3$ per voxel, and a flip angle=$12°$. The MRI-T1 and MRI-T2 studies are composed by 160 and 45 slices, respectively. The acquisition was performed using a parallel imaging factor of 2.

3.1.2 dMRI dataset

dMRI data of the head were acquired from a healthy subject on a General Electric Signa HDxt 3.0T MR scanner using the body coil for excitation, and an 8-channel quadrature brain coil for reception. We employ 60 gradient directions with a value for b equal to $1000\ S/mm^2$. The study contains $128 \times 128 \times 33$ images in axial plane.

Also, we downloaded a full study from the *human connectome project*: `https://www.humanconnectome.org/`, specifically from *MGH Adult Diffusion Data* repository. The data were collected from a male subject (age between 20-24) on the customized Siemens 3T Connectom scanner, a 64-channel tight-fitting brain array coil was used for data acquisition, and 128 gradient directions with a value for b equal to 5000 S/mm^2. The study contains $140 \times 140 \times 96$ images in the axial plane with isotropic voxel size of 1.5 mm, and we select a region of interest (ROI) of $40 \times 40 \times 10$ voxels centered in the corpus callosum.

3.2 Methodology

We propose the following set of activities for each specific aim, in order to reach our general aim.

SPECIFIC AIM 1. *To develop a stochastic methodology for super-resolution of anatomical MRI (T1 y T2) through scalar stochastic processes for supporting tissue segmentation.*

We will develop a supervised learning methodology based on Gaussian processes (GP) for super-resolution in anatomical MRI: T1 and T2 (See activities 1.1,

1.2, 1.3). The validation is performed as follows: First, using a dataset, we downsample the T1 and T2 images to obtain 256×256 multi-slice low-resolution (LR) images. Next, we perform the patch-based GP algorithm to build the 512×512 multi-slice super-resolution (SR) images. We compare against two common methods based on nearest-neighbor (NN) interpolation [46], and B-splines [20] to estimate 512×512 high-resolution (HR) images starting from the LR studies. Also, we compare with the GP methodology proposed in [75]. We apply a 2D filtering stage to improve edge enhancement for both methods based on GPR. Next, we compare all the HR and SR images with respect to the ground truth. As the T1 and T2 studies are 3D images, we evaluate the performance of the methods by computing the average of the mean squared error (MSE) obtained in each slice respect to the ground truth slices. Finally, using the MRI-T1 study, we validate morphologically the SR algorithms through segmentation of different brain tissues such as cerebrospinal fluid (CSF), white matter (WM) and gray matter (GM). We perform the morphological validation using the segmentation module with atlas of the multi-platform software package for visualization and medical image computing (3D-Slicer), available at `http://www.slicer.org/` [76]. We average the MSE obtained in each slice per each segmented tissue.

- **Activity 1.1:** To build a database of T1 and T2 images with studies acquired in the *laboratorio de procesamiento de imagen médica* in Madrid-Spain.
- **Activity 1.2:** To code the model based on GPs in Matlab®.
- **Activity 1.3:** To validate the proposed methodology employing error analysis and tissue segmentation metrics. We compare with gold standards images.

SPECIFIC AIM 2. *To build a stochastic interpolation approach applied in rank-2 tensor fields for enhancing spatial resolution in diffusion tensor images (DTI) and preserving fractional anisotropy of diffusion images. Application to fiber tracts reconstruction.*

We propose two different stochastic modeling of DTI:

First, We assume that a DTI field follows a generalized Wishart process (GWP). A GWP is a collection of symmetric positive definite random matrices indexed by an arbitrary dependent variable [65], i.e. the $[x, y, z]^\top$ position. In this context, we use it to model the entire DTI field $D(x, y, z)$. Then, through approximate Bayesian inference (i.e Elliptical slice sampling and Markov Chain Monte Carlo methods), we estimate the optimal parameters of the model.

Second, we propose a feature-based scheme combined with a stochastic treatment for interpolation of DT fields. Similar to [23], we decompose the tensors in eigenvalues (three) and Euler angles (three) for obtaining six features. Then, we interpolate non-linearly each feature employing a multi-output Gaussian process (MOGP)[77]. In this context, we consider each feature as an output of the MOGP indexed in space coordinates (x, y, z). The MOGP has a common covariance. For this reason, the outputs are correlated, unlike to [23], where the features are interpolated linearly and independently. We test toy and real datasets.

Stochastic modeling of DTI fields has some advantages: positive definite matrices, robustness to noise, smooth transition among nearby tensors and good accuracy for estimating new data. For both methods, the validation is as follows: we downsample the dataset in a factor of two. Original DT fields (without downsampling) are our gold standards or ground truth data, and we train the GWP with the full tensors or the MOGP with features extracted from downsampled fields. We compare our

methodology with approaches presented in [19, 23]. We apply two metrics: Frobenius norm and Riemann distance for evaluating accuracy in interpolation, and we test the fractional anisotropy (FA) curve for evaluating properties of diffusion tensors. Finally, we evaluate quality of tissue segmentation (GM, WM and CSF) from FA maps.

As ground truth (gold standard) we employ three different types of data. The first one corresponds to a synthetic DTI field. The second one corresponds to a simulation of crossing fibers using the algorithm of the fanDTasia toolBox [28], available at http://www.cise.ufl.edu/~abarmpou/lab/fanDTasia/. The third one, corresponds to a DTI dataset estimated from real dMRI through the RESTORE method [78]. We use two metrics to measure the differences between the interpolated fields and the ground truth, the Frobenius norm, and the Riemman distance, defined by

$$\mathrm{Frob}(T_1, T_2) = \sqrt{\mathrm{trace}\left[(T_1 - T_2)^\top (T_1 - T_2)\right]},$$

$$\mathrm{Riem}(T_1, T_2) = \sqrt{\mathrm{trace}\left[\log(T_1^{-1/2} T_2 T_1^{-1/2})^\top \log(T_1^{-1/2} T_2 T_1^{-1/2})\right]},$$

where T_1 and T_2 are the estimated and the ground truth tensors, respectively. The error metrics are computed for each *voxel*. We report the mean and standard deviation for the errors over the predicted data.

- **Activity 2.1:** To build a dMRI data base with studies recorded in *laboratorio de procesamiento de imagen médica* in Madrid-Spain.
- **Activity 2.2:** To perform Monte carlo methods for estimating optimal parameters in the GWP and to code the model in Matlab®.
- **Activity 2.3:** To validate the model developed in activity 2.2 over synthetic and real dMRI rank-2 data employing error analysis and fractional anisotropy.

SPECIFIC AIM 3. *To formulate a probabilistic approach for interpolation of higher order tensors (HOT) fields allowing the enhancement of spatial resolution for construction of generalized anisotropy maps.*

To the best of our knowledge, there is not a generalized methodology for HOT fields interpolation (no matter the rank), that retains all mandatory constraints for tensorial representation of dMRI. For this aim, we will introduce methodologies to perform interpolation in HOT fields of any order. As first approximation, We will formulate new stochastic processes over tensors based on the Tucker decomposition (TDP) and the canonical decomposition (CPD).

We test the TDP and CDP in HOT fields of rank 2,4 and 6 in three different types of datasets. First, we obtain synthetic HOT fields from a random generative model. Second, we obtain synthetic crossing fibers fields obtained from the algorithm of the fanDTasia toolBox [28], available at http://www.cise.ufl.edu/~abarmpou/lab/fanDTasia/. Finally, we estimate HOT data from a real dMRI study using the method proposed in [28]. The dMRI study was obtained from a human brain of a healthy subject on a General Electric Signa HDxt 3.0T MR scanner, 8-channel quadrature brain coil for reception, and 90 gradient directions with a value for b equal to 1000 S/mm^2. The study contains $128 \times 128 \times 33$ images in axial plane. As ground-truth or gold standard we use the original HOT data (synthetic and real), then we downsample the HOT fields by a factor of two. The downsampled fields are the training sets. After we train the TDP and CDP, we compute the predictive distribution for the HOT fields. For rank-2 data, we compare our approach with direct

linear interpolation [18], log-Euclidean interpolation [19], and generalized Wishart processes (GWP) [40].

For rank-4 and 6, we compare against direct linear interpolation and raw dMRI interpolation with b-splines (only for the real dMRI data set). For a quantitative evaluation, we calculate an error metric based on the tensorial Frobenius distance (see eq. (2.43)) between the interpolated field and the respective ground-truth, evaluating only the predicted tensors. Also, we test morphological validation employing fractional anisotropy (FA) maps and tractography analysis (2D and 3D) for rank-2 tensors. FA is a measurement of anisotropy levels in dMRI, where a 0 value corresponds to an isotropic tensor and 1 refers to a full anisotropic tensor. In the case of 3D tractography, we evaluate the number of generated fibers and the average length of tracts. For rank-4 and 6, we evaluate generalized anisotropy (GA) curves. GA is a generalization of FA for higher orders. Figure 3.1 illustrates a flow diagram of the proposed framework.

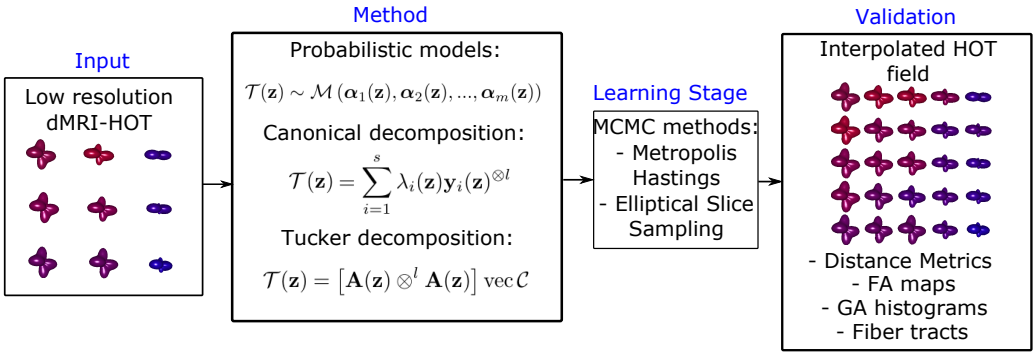


FIGURE 3.1: Flow diagram of the proposed framework. The input is a low resolution field (i.e. 3×3 tensors). Then, we model the HOT field with CDP or TDP. Next, we find the posterior of parameters employing MCMC methods. Finally, we validate the enhanced resolution field (i.e. 5×5 tensors) comparing with a gold standard through distance metrics, FA maps, GA histograms, and fiber tracts.

- **Activity 3.1:** To formulate the mathematical definition of proposed methods.
- **Activity 3.2:** To perform inference methods for estimating relevant parameters for both models formulated in activity 3.1.
- **Activity 3.3:** To code the models of activity 3.2 in Matlab®.
- **Activity 3.4:** To validate the proposed methodologies in toy and real data, using distance metrics and generalized anisotropy.

Chapter 4

Results and discussion

4.1 Resolution enhancement of structural MRI (T1-T2)

4.1.1 Validation with error metric

We make a direct comparison between the SR-MR images obtained for each method and the ground truth studies. As we pointed out before, we compute the average MSE for the whole up-sampled MRI studies (T1 and T2). Table 4.1 shows the average MSE results for B-splines, nearest-neighbor (NN), the GPR proposed in [75] (SR-GPR1), and our patch-based GPR methodology (SR-GPR2). In figure 4.1, we show graphic errors obtained in SR-MR images validation (T1 and T2). The ground truth for T1 slice-100 and T2 slice-20 are showed in (a) and (b), respectively. Sub-figures (c),(d),(e) and (f) correspond to absolute error images of the SR-T1 image for B-spline, nearest-neighbor (NN), SR-GPR1, and SR-GPR2, respectively. Finally, in sub-figures (g),(h),(i) and (j), the absolute error images for SR-T2 image are showed in the same order used for SR-T1 results.

TABLE 4.1: Average MSE results obtained for T1 and T2 respect to the MRI ground truth

MR study	B-spline $\mu \pm \sigma$	NN $\mu \pm \sigma$	SR-GPR1 $\mu \pm \sigma$	SR-GPR2 $\mu \pm \sigma$
T1	27.95 ± 7.282	38.12 ± 10.45	55.89 ± 20.47	$\mathbf{7.806 \pm 1.552}$
T2	48.83 ± 13.14	60.44 ± 16.05	72.99 ± 22.12	$\mathbf{25.65 \pm 7.659}$

4.1.2 Morphological validation

Understanding the importance of tissue segmentation for clinical procedures and diagnosis, we validate morphologically the SR methods over the MRI-T1 study. For this test, the gold standards are the probabilistic maps of segmentations achieved with the ground truth data using 3D slicer [76]. We segment three types of tissues such as cerebrospinal fluid (CSF), white matter (WM), and gray matter (GM). Similarly, we compute the average MSE in all the tissue segmentations obtained from the SR-T1 results. Table 4.2 shows the average MSE results for B-splines, nearest-neighbor (NN), SR-GPR1 and SR-GPR2. Also, in Figure 4.2, we show some graphic errors obtained for morphological validation. Using the T1 slice-100 from Figure 4.1, the gold standards of CSF, WM and GM segmentation are showed in (a),(b) and (c), respectively. Sub-figures (d),(e),(f) and (g) correspond to absolute error images of CSF segmentation for B-splines, nearest-neighbor, SR-GPR1 and SR-GPR2, respectively. Following the same order, from sub-figure (h) to (k) and from sub-figure (l)

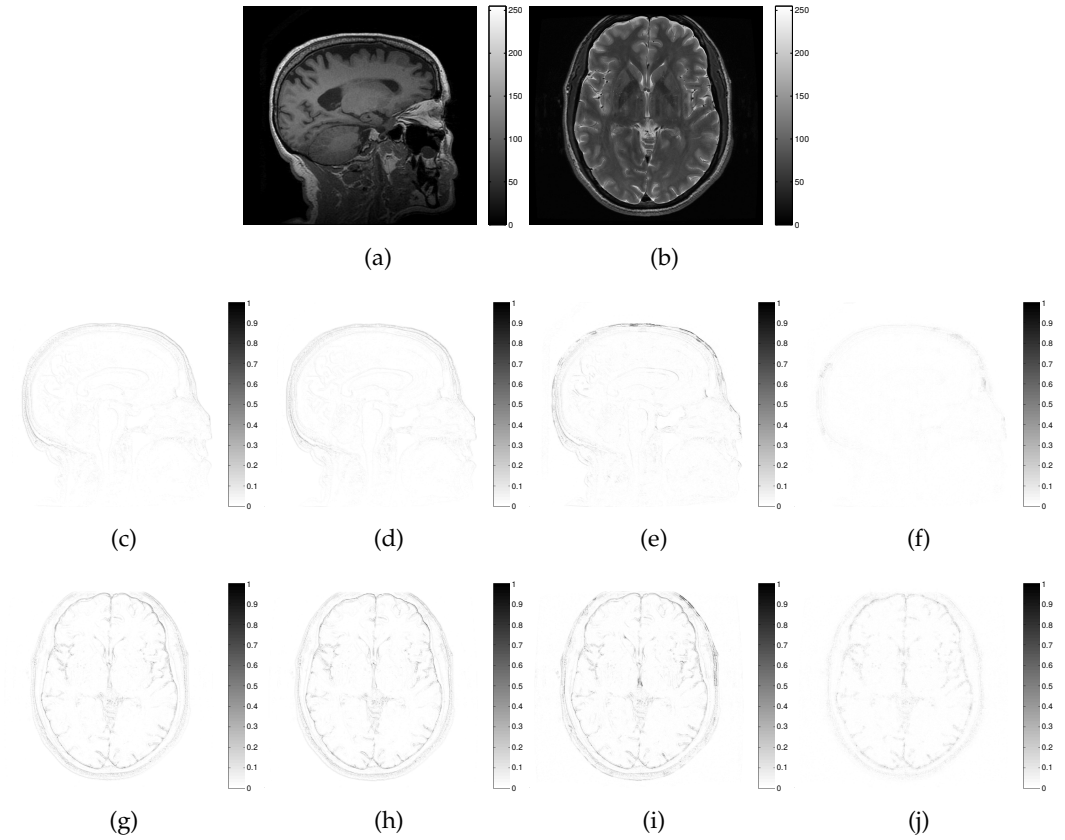


FIGURE 4.1: Ground truth and graphic errors for super-resolution (SR) MRI images validation. The ground truth for T1 slice-100 and T2 slice-20 are showed in (a) and (b), respectively. (c),(d),(e) and (f) correspond to absolute error images of SR-T1 studies for B-spline, nearest-neighbor, SR-GPR1 and SR-GPR2, respectively. Finally, in (g),(h),(i) and (j), the absolute error images for SR-T2 study are showed in the same order than SR-T1 results. The color-bars show the magnitude of the absolute error produced.

TABLE 4.2: Average MSE results obtained respect to the gold standard tissue segmentations.

Tissue	B-splines $\mu \pm \sigma$	NN $\mu \pm \sigma$	SR-GPR1 $\mu \pm \sigma$	SR-GPR2 $\mu \pm \sigma$
CSF	0.0154 ± 0.0142	0.0148 ± 0.0140	0.0079 ± 0.0064	$\mathbf{0.0044 \pm 0.0040}$
WM	0.0362 ± 0.0297	0.0357 ± 0.0322	0.0166 ± 0.0124	$\mathbf{0.0084 \pm 0.0065}$
GM	0.0422 ± 0.0300	0.0406 ± 0.0312	0.0203 ± 0.0132	$\mathbf{0.0105 \pm 0.0071}$

to (o), the absolute error images for WM and GM segmentation are showed, respectively.

4.1.3 Discussion

From Table 4.1, we can observe that under the same experimental conditions, our methodology for super-resolution (SR-GPR2) outperforms both MRI-T1 and MRI-T2 validation studies. Although the T2 study has a lower signal noise ratio (SNR) than the T1 study, SR-GPR2 works better than the other methods. Regardless of the type of MRI study, our proposed method based on GPR achieves a higher accuracy in super-resolution. The low performance obtained by SR-GPR1 is due to the natural smooth behavior of GPs. When the SR-GPR1 estimates new pixels, the GP tends to generate a blurring effect over edges in the up-sampled MR images. For this reason, the low capability to capture high contrast transitions may be a considerable drawback in SR-GPR approaches. In SR-GPR2, this problem is partially solved with the 2D filtering stage. However, for SR-GPR1 this issue remains. When, we apply the filtering to the comparison methods, the outcomes are even worse than those obtained without filtering. From Figure 4.1, we observe that high error values are produced by strong changes in edge regions (e.g. skull, lobes and brain structures), concluding that B-splines, NN and SR-GPR1 fail for performing super-resolution in high contrast areas. It is clear that error images for T1 and T2 show a higher accuracy for our proposed method (SR-GPR2). But, there are some prediction mistakes in several pixels due to changes of high contrast that SR-GPR2 cannot avoid. Perhaps, a more robust filtering stage could improve the pixels estimation at edges.

Morphological validation is a key procedure to establish if a super-resolution approach is appropriate or not. Brain tissue segmentation is one of the main applications of MRI-T1 studies. In surgical planning of neuro-degenerative diseases, anatomical segmentation is necessary in order to identify types of tissue in MRI studies. However, LR images present problems such as partial volume effect [12]. This leads to erroneous segmentations because there is not a significant difference among the brain tissue types. Therefore, SR becomes a necessary task for achieving a refined and accurate segmentation.

From Table 4.2 and Figure 4.1, we can observe that our methodology also outperforms the results in morphological validation. Although SR-GPR1 already outperforms the segmentation provided by the standard techniques, for all the different tissues, SR-GPR2 works better due its ability to represent high contrast transitions. We conclude then that the patch-based GPR with 2D post-filtering stage provides promising result in SR-MR images.

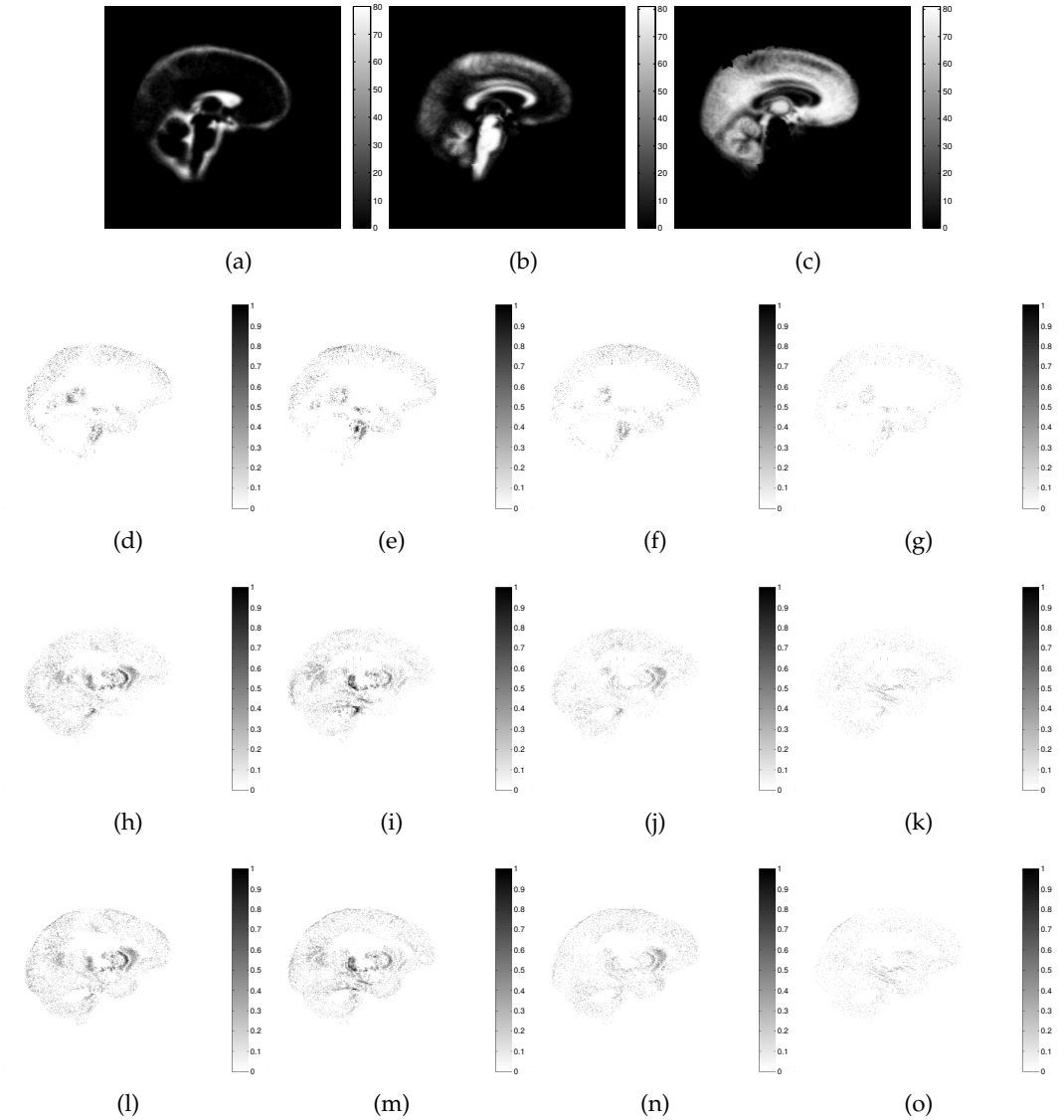


FIGURE 4.2: Ground truth and graphic errors for morphological validation. Using the T1 slice-100 from Figure 4.1, the gold standard of CSF, WM and GM segmentation are showed in (a),(b) and (c), respectively. (d),(e),(f) and (g) correspond to absolute error images of CSF segmentation for B-splines, nearest-neighbor, SR-GPR1 and SR-GPR2, respectively. Following the same order, from sub-figure (h) to (k) and from sub-figure (l) to (o), the absolute error images for WM and GM segmentation are showed, respectively. The color-bars show the magnitude of the absolute error produced.

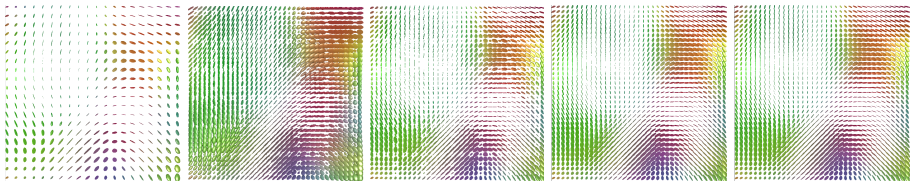
4.2 DTI (Rank-2 tensors) interpolation

4.2.1 Generalized Wishart Processes

We compare the proposed generalized Wishart process with linear [18] and log-euclidean interpolation [19].

Synthetic Data

We generate noisy random DTI data to construct a $2D$ field of 37×37 tensors. We assume 25 gradient directions for generating DTs, and b value of $1000\ s/mm^2$. In Figure 4.3 we can see the initial downsampled DTI field, linear and log-euclidean interpolation, the interpolated field with GWP, and the ground truth respectively. Table 4.3 shows the error metrics.



(a) (b) (c) (d) (e)

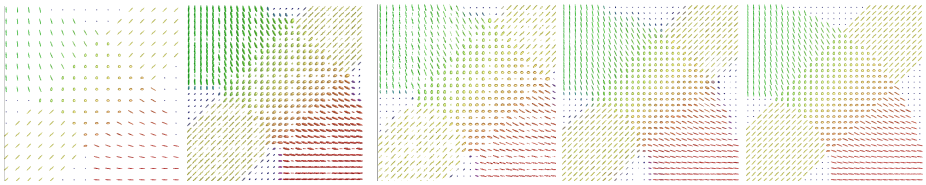
FIGURE 4.3: Graphic results for DTI interpolation $(2\times)$ applied in synthetic data. (a) Downsampled DTI field (data used for estimation). (b) Linear interpolation. (c) Log-euclidean interpolation. (d) Interpolation with GWP. (e) Ground truth.

TABLE 4.3: Metric results for synthetic DTI field

	Frobenius distance $(\times 10^{-5})$	Riemman distance
GWP	7.06 ± 1.51	0.160 ± 0.125
linear interpolation	50.11 ± 4.26	8.54 ± 1.36
log-euclidean	35.25 ± 3.92	6.34 ± 1.22

DTI from crossing fibers

One of the most critical DTI datasets correspond to crossing fibers. We generate this type of DTI field through FanDTasia toolbox [28]. This dataset describes a $2D$ crossing fiber field with 31×31 tensors. Figure 4.4 and Table 4.6 show the comparative results.



(a) (b) (c) (d) (e)

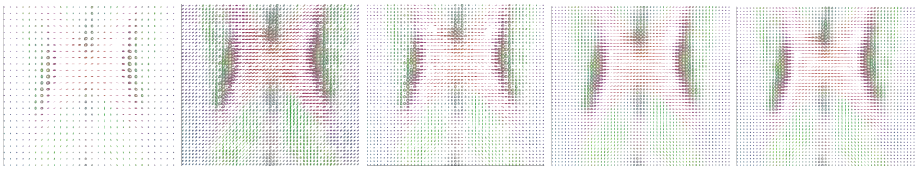
FIGURE 4.4: Graphic results for DTI interpolation $(2\times)$ applied in crossing fibers field. (a) Downsampled DTI field (data used for estimation). (b) Linear interpolation. (c) Log-euclidean interpolation. (d) Interpolation with GWP. (e) Ground truth.

TABLE 4.4: Error measures for crossing fibers in a DTI field.

	Frobenius distance $(\times 10^{-5})$	Riemman distance
GWP	18.11 ± 11.82	0.184 ± 0.114
linear interpolation	73.12 ± 8.26	11.14 ± 2.65
log-euclidean	61.09 ± 6.15	9.74 ± 1.67

Real DTI field estimated from dMRI

Finally, we test our method in real DTI data estimated from dMRI acquired in a human subject. The field corresponds to an axial slice with 49×55 tensors. Figure 4.8 and Table 4.7 show the comparative results.



(a) (b) (c) (d) (e)

FIGURE 4.5: Graphic results for DTI interpolation $(2\times)$ applied in real DTI data. (a) Downsampled DTI field (data used for estimation). (b) Linear interpolation. (c) Log-euclidean interpolation. (d) Interpolation with GWP. (e) Ground truth.

TABLE 4.5: Error measures for the real DTI field example

	Frobenius distance $(\times 10^{-5})$	Riemman distance
GWP	6.26 ± 3.20	0.146 ± 0.080
linear interpolation	45.76 ± 7.21	7.25 ± 2.10
log-euclidean	31.67 ± 6.10	6.89 ± 1.86

Discussion

Linear and log-euclidean interpolation seek to minimize geodesic distances. The geometric (Riemann and Euclidean) approaches work well in smooth DTI fields. However, they reduce their performance in presence of high level of noise. For example, when we interpolate the synthetic noisy data (Figure 4.3 and Table 4.3), we can observe a swelling effect for the estimated tensors in the new input locations, when using linear interpolation. This is a critical issue, because this effect modifies the fractional anisotropy maps. Another big problem with linear interpolation is the possibility of obtaining non-positive definite tensors. Although, non-positive definite tensors are avoided with log-euclidean interpolation, the accuracy of tensor estimation in new input locations is not satisfactory. On the other hand, GWP guarantees positive definite tensors because of its mathematical construction. Also, this probabilistic model is more robust to noise, and it can keep the smooth transition among spatially nearby data. This property avoids the swelling effect. If we look at the results for more complex data like crossing fibers DTI, and the real DTI field (Figures 4.4,4.8 and Tables 4.6,4.7), we can see better accuracy results for GWP. Both average distances (Frobenius and Riemann) are smaller in GWP than Linear and Log-Euclidean methods. Similarly, graphic results of DTI fields are smoother

for GWP interpolation. The GWP takes into account the global spatial behavior of the DTI data, while geometric approaches estimate tensors only with the nearest tensors.

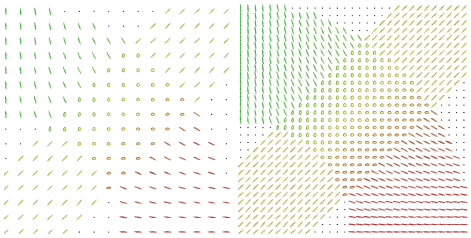
A drawback in GWP happens when there are strong changes in tensor orientation (i.e. crossing fibers). The GWP cannot capture extreme modifications in data of crossing fibers. Recall that a GWP is a superposition of Gaussian processes (GP) and GPs are modeled with smooth kernels functions. The best alternative to model crossing fibers is through Higher Order Tensors (HOT). Nevertheless a GWP does not describe HOT. However, for the geometric approaches, strong changes in tensor orientation generate worse results than GWP.

In summary, the probabilistic modeling of DTI fields that we employ here allows a better description of global spatial transition in tensorial imaging. Geometric approaches are fine for simple DTI fields. However in real applications, DTI data are very complex (High level of noise, Heterogeneous data, Non-positive definite tensors, etc). The GWP has many advantages, for example: it guarantees positive definite tensors, robustness to noisy data, smooth transition among nearby data, no swelling effect, it keeps important properties of DTI (FA maps) and excellent accuracy.

4.2.2 Multi-Output Gaussian Processes

Crossing fibers dataset

It is a simulated dataset of a $2D$ crossing fiber field with 31×31 tensors (see figure 4.6). Table 4.6 show the comparative errors for interpolation. MOGP refers to multi-output Gaussian processes, GWP is the generalized Wishart process [40], FBL is the feature-based approach with linear interpolation [23] and log-euclidean is the scheme presented at [19]. Figure 4.7 shows the fractional anisotropy (FA) curve of each method.



(a) (b)

FIGURE 4.6: Crossing fibers field. (a) Downsampled DT field (data used for training). (b) Original dataset (Ground truth).

TABLE 4.6: Error measures for crossing fibers in a DTI field.

	Frobenius distance ($\times 10^{-5}$)	Riemman distance
MOGP	17.45 ± 9.12	0.195 ± 0.101
GWP	18.11 ± 11.82	0.184 ± 0.114
FBL	30.12 ± 8.26	2.67 ± 0.551
log-euclidean	61.09 ± 6.15	9.74 ± 1.67

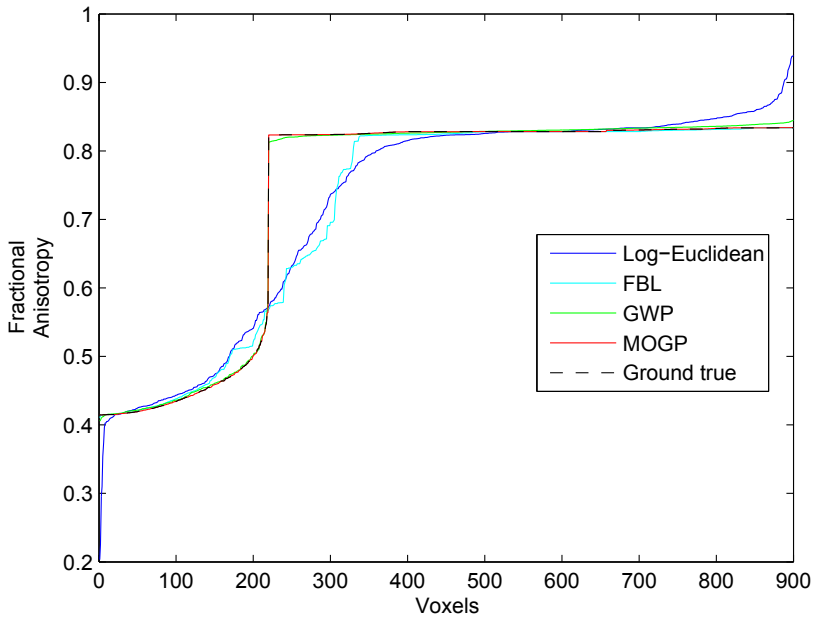
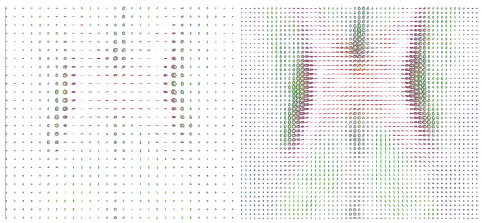


FIGURE 4.7: Fractional anisotropy (FA) curve for crossing fibers after interpolation

Real DT field

The field corresponds to an axial slice with 49×55 tensors (See figure 4.8) acquired from a human subject. Table 4.7 show the comparative errors and figure 4.9 shows the FA curve. MOGP, GWP, FBL and log-euclidean were defined in subsection 4.2.2



(a) (b)

FIGURE 4.8: Real DT field. (a) Downsampled DT field (data used for training). b) Original dataset (Ground truth).

TABLE 4.7: Error measures for the real DT field example

	Frobenius distance ($\times 10^{-5}$)	Riemman distance
MOGP	5.94 ± 2.15	0.151 ± 0.043
GWP	6.26 ± 3.08	0.146 ± 0.080
FBL	18.43 ± 4.10	4.56 ± 1.18
log-euclidean	31.67 ± 6.10	6.89 ± 1.86

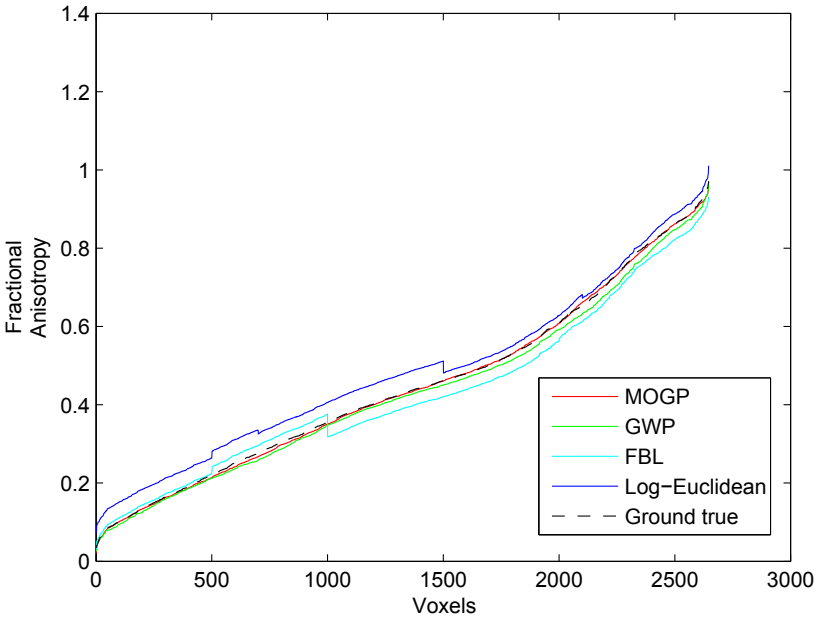


FIGURE 4.9: fractional anisotropy (FA) curve for the real DT field after interpolation

Discussion

If we observe all results from experiments on both databases, it is clear that probabilistic methods (MOGP and GWP) exhibit a better performance than traditional approaches (log-euclidean and FBL). This can be explained because a known advantage of probabilistic methods is robustness to presence of noise in data. Also, non-linear regression with Gaussian processes has proved to be suitable for all type of data, even for small datasets. Outcomes of MOGP and GWP are not statistically different, for this reason we cannot establish which one is better. When we observe the quantitative evaluation of methods (tables 4.6 and 4.7), the difference between MOGP and log-euclidean or FBL is very considerable, no matter the metric distance. While, a direct comparison among methods does not ensure that interpolation can preserve the properties of the original tensor field, the analysis of FA curves allows to guarantee that MOGP does not neglect any constraint of DTs. Figures 4.7 and 4.9 show the evolution of FA across the fields for all methods. Again, the FA curve for MOGP follows the ground-true tendency as close as possible compared to the other methods. The FA analysis is very relevant, due to it confirms that important properties of diffusion tensors are preserved with MOGP.

Although, FBL [23] is a feature-based method that decomposes the tensors in a similar way to our proposed methodology, a deeper insight shows that our scheme with MOGP captures the correlation among extracted features. In this context, we consider the six features as outputs of a Gaussian process, where spatial evolution of an output affects the other five. This additional information is a key factor for a better learning instead of interpolating features independently. A drawback of MOGP appears when the DT field has strong changes among nearby tensors (i.e crossing fibers). This issue is due to the soft nature of kernels that defines the covariances matrices inside GPs. In these cases, the accuracy of prediction is affected. Finally, we think that a probabilistic treatment of tensorial data gives some important advantages for processing DT fields such as: high accuracy, robustness, and preservation of properties and constraints of DTs.

4.2.3 Multi-Output Gaussian Processes with non-stationary kernel

Crossing fibers

Crossing fibers data is one of most difficult field to interpolate, because transition among tensors is very abrupt. We test the NmoGp over a 2D crossing fibers field of 22×22. We use five latent functions $(r = 5)$ and we employ a rational quadratic kernel. The figure 4.10 (a) shows the crossing fibers field, figure 4.10(b) is the down-sampled tensor field (used for training), and figure 4.10(c) illustrates the validation tensors. Also, we show Riemann error maps in figures 4.10(d)(e)(f). If we observe in detail, error values are very low over smooth regions of the tensor field for all methods (blue zones in error maps of fig. 4.10). However, when we evaluate strong transition regions, interpolation is not straightforward, and the error is higher in this areas. Specifically, over this critic region of the field, our model achieves more precision than comparison methods. We can explain this because the non-stationary kernel can capture the dynamic of the tensor field. While it is true that there are some errors when we interpolate new tensors in the crossing fibers region, we can say that NmoGp outperforms to moGp and the linear method, when interpolation is challenging due to strong changes in neighboring tensors.

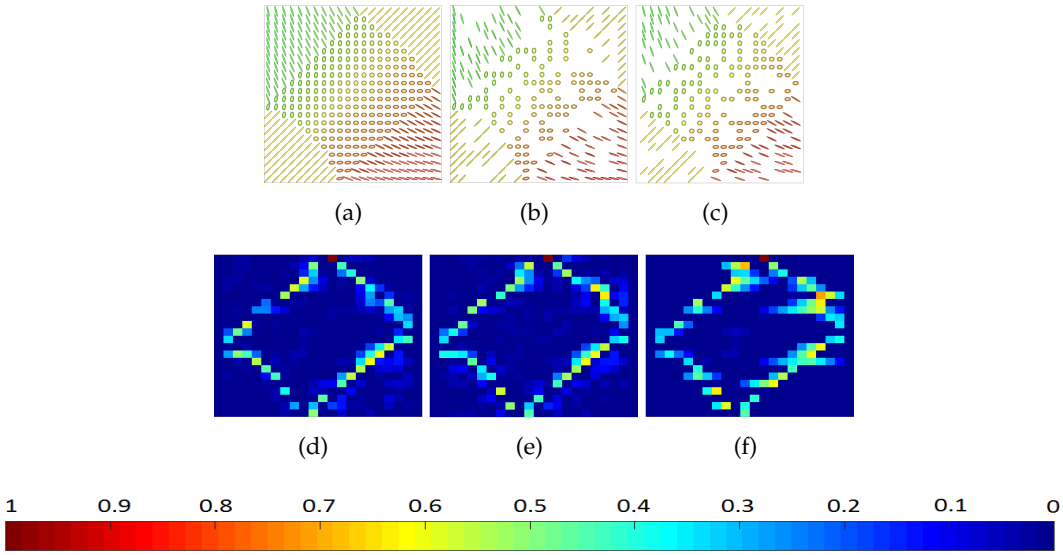


FIGURE 4.10: Interpolation over crossing fiber fields, (a) gold standard, (b) training field, (c) test field. (d),(e),(f) Riemann norm error maps, for the NmoGp, moGp and Linear model, respectively.

TABLE 4.8: Frobenius and Riemann distances for crossing fibers of NmoGp, moGp, and Linear methods

Model	Riem			Confidence Interval	Frob (10e-4)			Confidence Interval
NmoGp	0.1917	$\pm$	0.2772	[0.1561, 0.2273]	17.342	$\pm$	25.191	[1.4103, 2.0581]
moGp	0.2133	$\pm$	0.2839	[0.1768, 0.2498]	19.589	$\pm$	26.136	[1.6229, 2.2949]
Linear	0.2666	$\pm$	0.3624	[0.2200, 0.3132]	24.228	$\pm$	33.375	[1.9937, 2.8519]

Also, according to results in Table 4.8, the linear model obtained the higher error (Frobenius and Riemann). Linear model has a considerable lack because it does not consider correlation among the six features extracted from tensor decomposition.

This approach only interpolates linearly and separately each feature. On the other hand, moGp and NmoGp interpolate the six features simultaneously. The idea is to share a correlation among features. For this reason, there is an additional information allowing a better estimation of new data. The main difference between moGp and NmoGP is the type of kernel for constructing the covariance matrix inside the model. Thereby, moGp works with a single kernel (stationary) while our NmoGp is based on a non-stationary kernel. We see in Table 4.8, that global error of the NmoGp approach is lower than the moGp and linear method.

Real DT Field

We test the methods over a 2D diffusion tensor field obtained from real dMRI. The data corresponds to an axial slice with 40×40 tensors of the head. Figures 4.11 (a),(b),(c) correspond to the gold standard, training and test data respectively. Figures 4.11 (d),(e),(f) show the Riemann error maps. Again, Table 4.9 show the error distance for all tested methods and the confidence interval. A real dMRI tensor field is noisy and very heterogeneous. Therefore, simple interpolation methods fail to achieve a good accuracy. Also, probabilistic methods such as moGp and NmoGp have the robustness property. Again, NmoGp improves to the comparison methods according to outcomes of Table 4.9. This result is very relevant, because it confirms that proposed method can interpolate tensorial data, no matter the type of dataset. Finally, we can establish that insertion of a non-stationary kernel in multi-output Gaussian process increases its performance significantly.

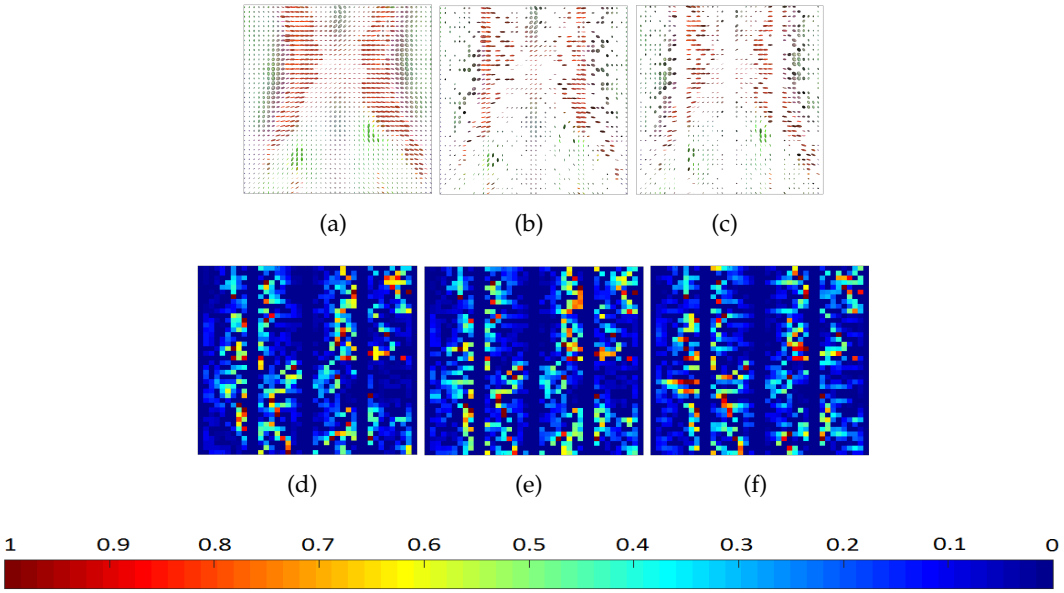


FIGURE 4.11: Interpolation over a 2D real crossing fiber field, (a) gold standard, (b) training field, (c) test field. (d),(e),(f) Riemann norm error maps for the NmoGp, moGp and Linear model, respectively.

TABLE 4.9: Frobenius and Riemann distances for real dMRI data of NmoGp, moGp, and Linear method

Model	Riem			Confidence Interval	Frob			Confidence Interval
NmoGp	11.222	$\pm$	1.041	[1.0512, 1.1932]	0.7254	$\pm$	0.5538	[0.6862 , 0.7646]
moGp	11.458	$\pm$	0.9588	[1.0780, 1.2136]	0.7450	$\pm$	0.4976	[0.7098, 0.7802]
Linear	11.569	$\pm$	1.044	[1.0859, 1.2279]	0.7561	$\pm$	0.5043	[0.7204, 0.7918]

4.3 Higher Order Tensors interpolation

We first illustrate how parameters of the models are trained using Monte Carlo methods. Second, we perform a Rician noise analysis evaluating error in interpolation of 2D synthetic tensor fields (30×30 voxels) for various signal-noise ratio (SNR) values. Then, we show quantitative and qualitative results obtained in three different dMRI data: a synthetic dataset, a crossing fibers field and a real dMRI study acquired from a human subject. For all datasets, we interpolate HOT fields of rank 2,4 and 6.

4.3.1 Monte Carlo methods

Figure 4.12 shows the samples and posterior distributions of relevant parameters for CDP and TDP, when we train a rank-2 synthetic field: length-scale (θ), y_{11} element of CDP and A_{11} element for TDP respectively. In this case, initial values of mentioned parameters are sampled from the priors. Recall that prior of θ is a log-normal distribution with $\mu_\theta = 0$ and $\sigma_\theta = 0.001$. while, elements of vectors $\mathbf{y}_i$ and matrix $\mathbf{A}$ follow independent Gaussian process $\mathcal{GP}(\mathbf{0}, k(\mathbf{z}, \mathbf{z}))$. A closer look to figure 4.12 demonstrates a stable behavior of all parameters analyzed. For example, θ hyper-parameter has some strong jumps, but its distribution function is not highly disperse. A similar tendency have y_{11} and A_{11}, where they present a Gaussian distribution.

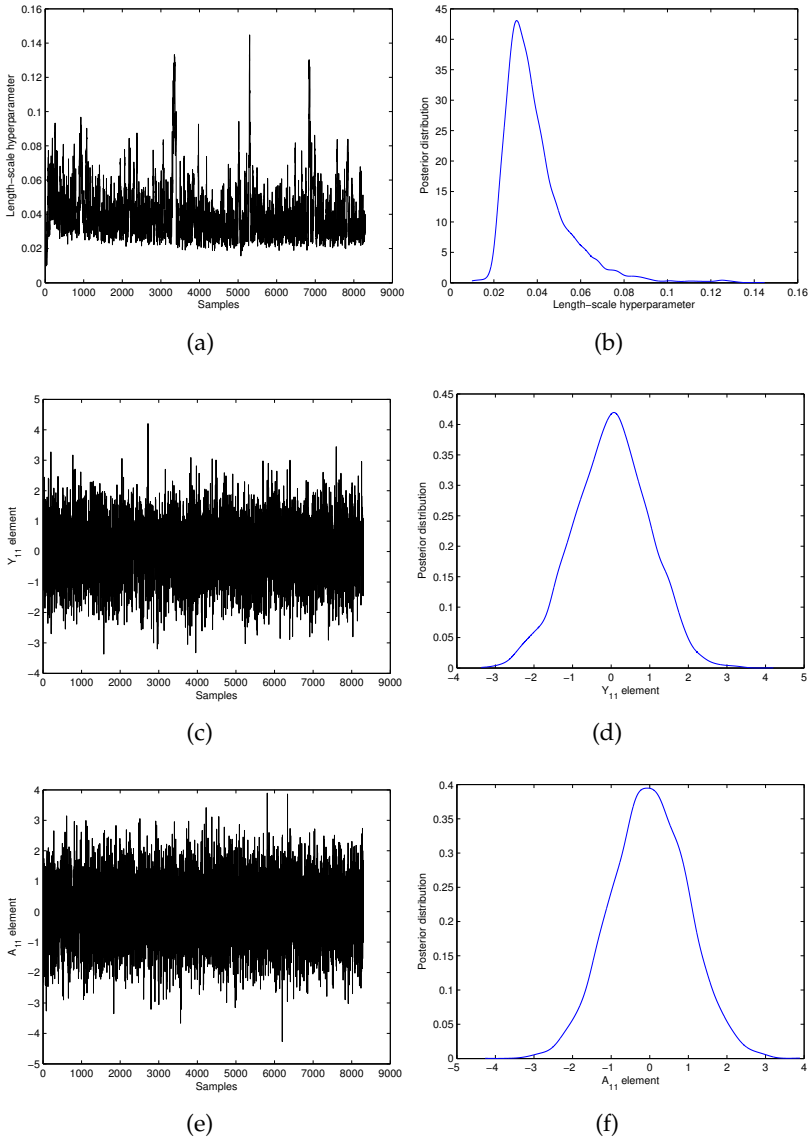


FIGURE 4.12: Samples and posterior distributions obtained for some relevant parameters: a) and b) correspond to the length-scale hyperparameter θ of the squared exponential kernel $k(\mathbf{z}, \mathbf{z}')$; c) and d) illustrate the element y_{11} of the unitary vector $\mathbf{y}_1$ used in CDP; e) and f) refer to the element A_{11} of the matrix $\mathbf{A}$ employed in TDP. Metropolis-Hastings method is utilized for θ and elliptical slice sampling is the algorithm used for elements of $\mathbf{y}_i$ and $\mathbf{A}$.

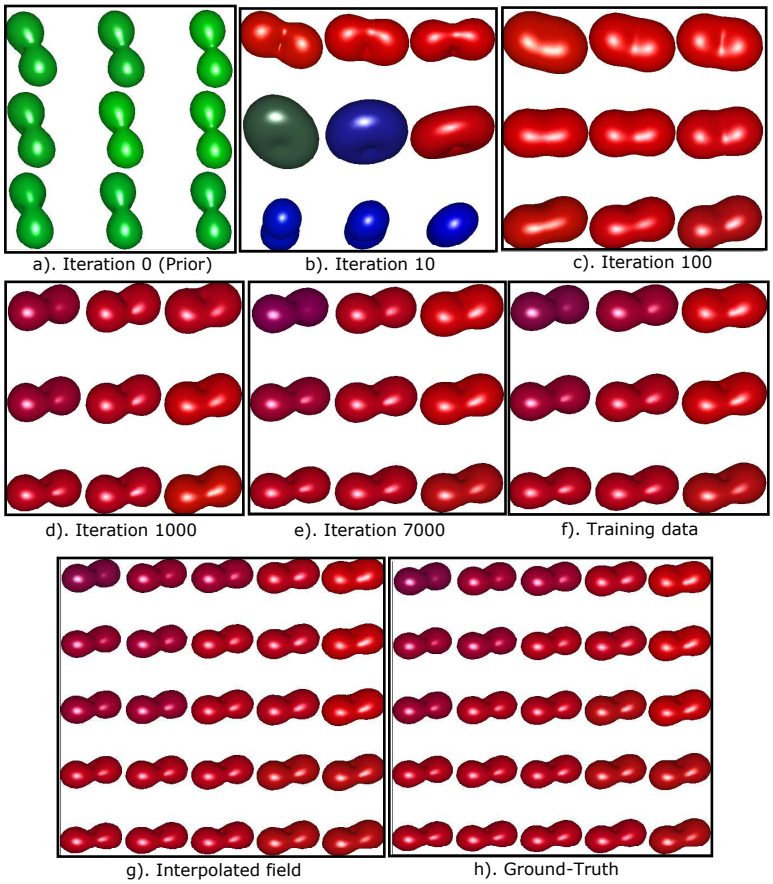


FIGURE 4.13: Learning and predictive process for a rank-2 synthetic field. Subfigures a) to e) show the evolution of a initial field obtained from sampling a prior (in this example we employ CDP, but a similar process occurs for TDP) until the learning stage is completed in iteration 7000; f) corresponds to the training data or low resolution field; g) corresponds to the interpolated or high resolution field, and h) is the ground-truth data.

Figure 4.13 describes the learning and prediction process for a rank-2 tensor field. As we can observe, the initial field obtained from the prior (CDP or TDP) is gradually modified until it achieves the values and shapes of the training data. To construct the posterior of CDP and TDP, we select the parameters with the biggest log-likelihood. Then, using the predictive distribution, we interpolate new data for enhancing spatial resolution of the tensor field.

4.3.2 Noise Analysis

We perform a noise analysis by testing the interpolation methods over tensor data corrupted with Rician noise for several SNR values. The noise is randomly distributed in the tensor field. Figures 4.14, 4.16, 4.18 show training data and ground truth for rank-2, 4 and 6 respectively. Also, figures 4.15, 4.17, and 4.19 show the mean error and standard deviation of interpolation in rank-2, 4 and 6 tensor fields. For all magnitude variations of SNR and different tensor orders, the proposed methods outperform to the comparison approaches.

Rank-2 data

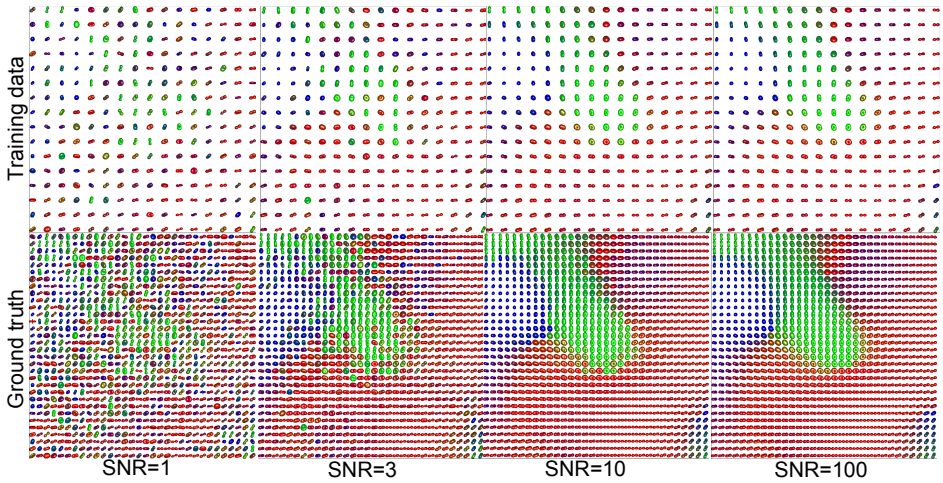


FIGURE 4.14: Rank-2 synthetic data corrupted with Rician noise for several SNR values. Superior and inferior rows correspond to the training data and ground truth respectively.

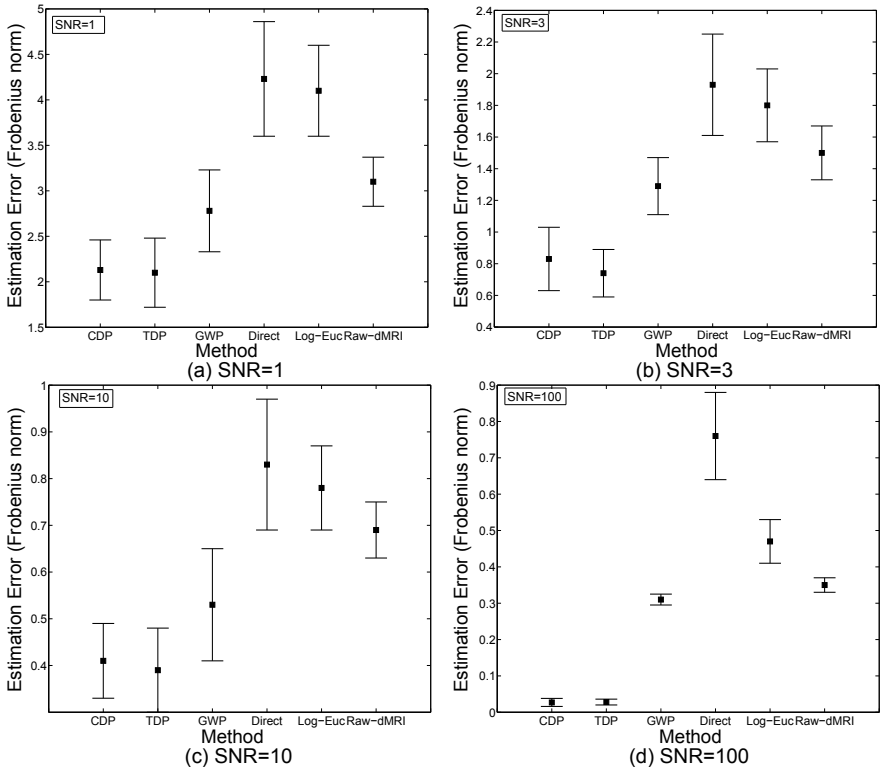


FIGURE 4.15: Mean and standard deviation of interpolation error (Frobenius norm) in rank-2 synthetic data for SNR equals to 1, 3, 10, and 100.

Rank-4 data

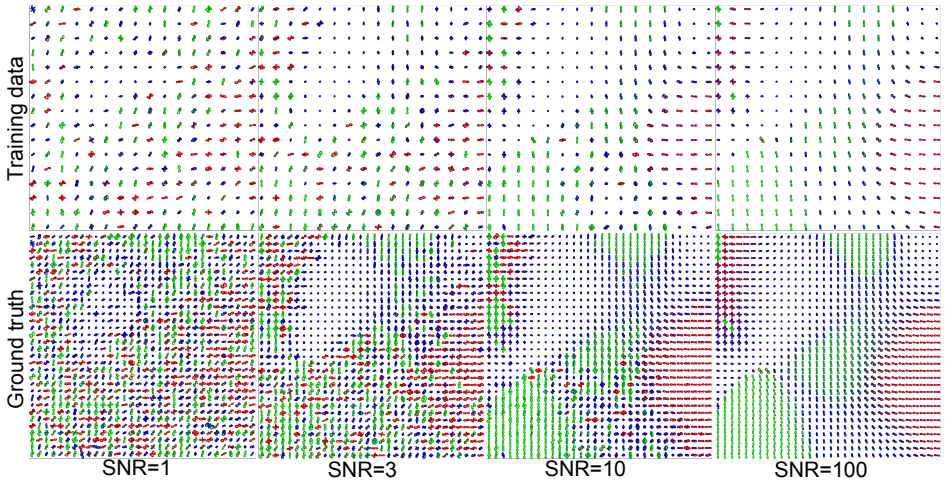


FIGURE 4.16: Rank-4 synthetic data corrupted with Rician noise for several SNR values.

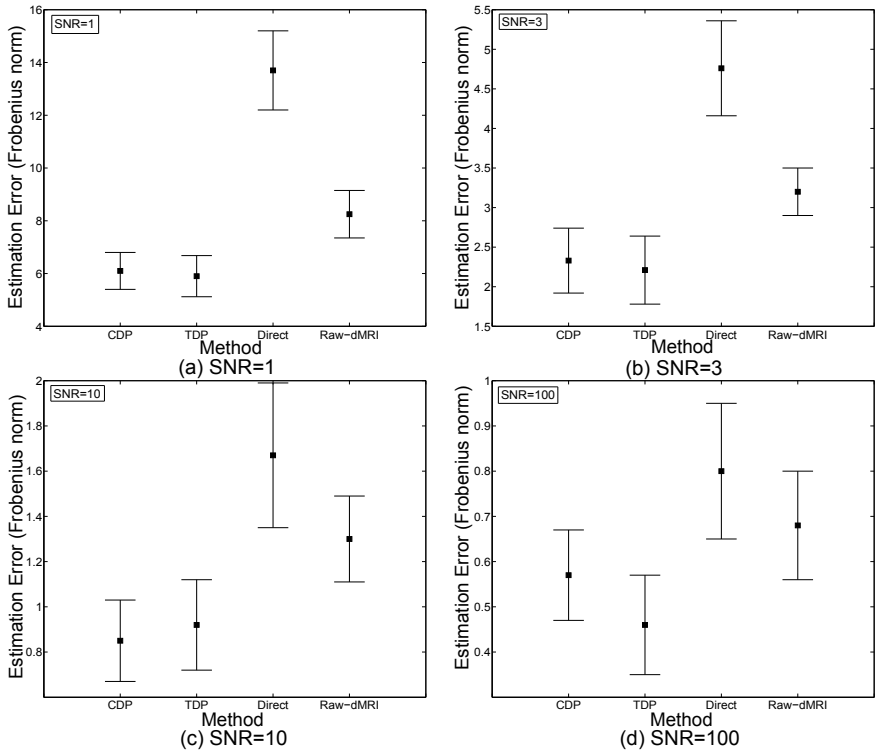


FIGURE 4.17: Mean and standard deviation of interpolation error in rank-4 synthetic data for several SNR values.

Rank-6 data

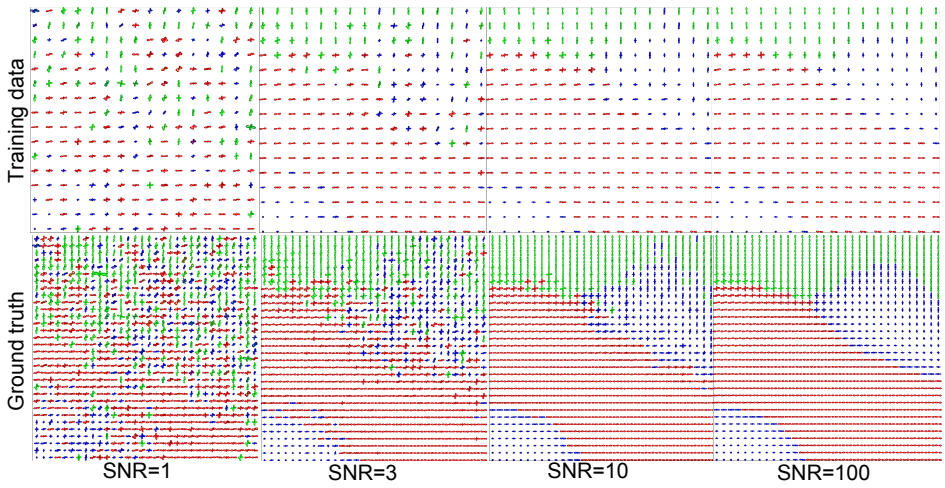


FIGURE 4.18: Rank-6 synthetic data corrupted with Rician noise for several SNR values.

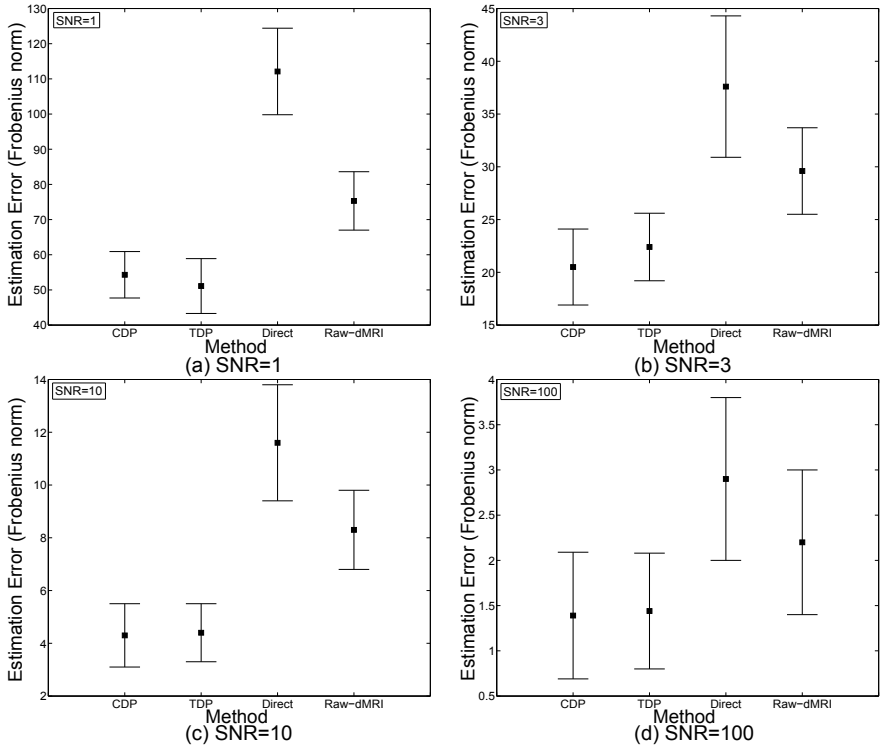


FIGURE 4.19: Mean and standard deviation of interpolation error in rank-6 synthetic data for several SNR values.

4.3.3 HOT fields interpolation in synthetic data

Rank-2 Results

Figure 4.20 and table 4.10 show results for the rank-2 synthetic data when we evaluate Frobenius distance and mean squared error (MSE) of Fractional Anisotropy (FA). Figure 4.21 show the morphological validation through FA maps and 2D fiber

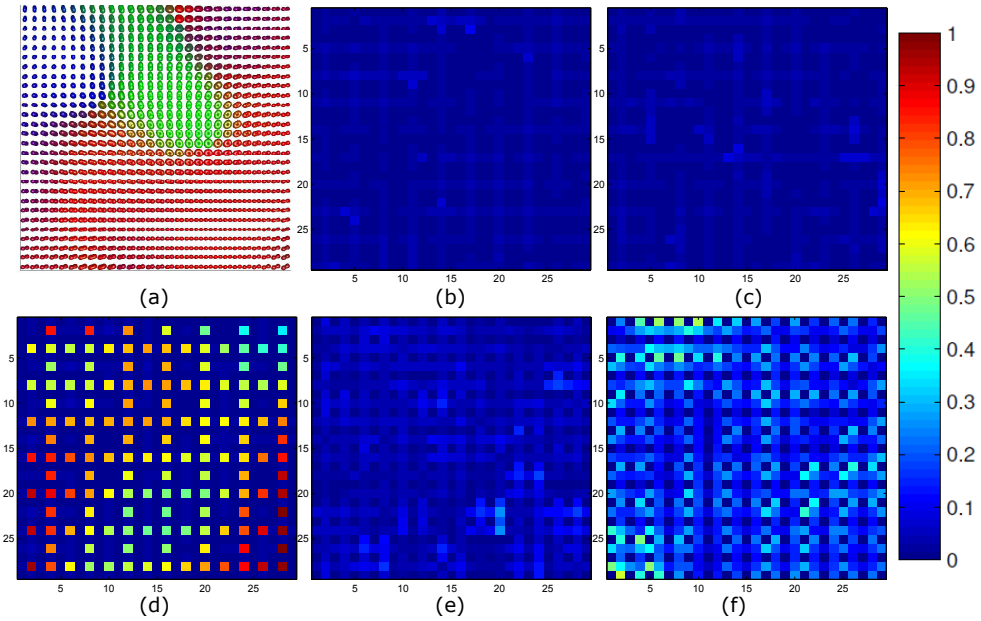


FIGURE 4.20: Normalized error maps for interpolation of a rank-2 synthetic HOT field: (a) Ground-truth data, (b) Canonical decomposition process (CDP), (c) Tucker decomposition process (TDP), (d) Direct interpolation , (e) Generalized Wishart processes (GWP), and (f) log-Euclidean.

TABLE 4.10: Frobenius distance and MSE of FA for rank-2 synthetic HOT field

	D. interpolation	Log-Euclidean	GWP	TDP	CDP
Frobenius distance.	0.756 ± 0.565	0.468 ± 0.312	0.313 ± 0.423	0.028 ± 0.032	0.027 ± 0.029
MSE of FA ($\times 10^{-5}$)	46.80 ± 19.00	23.64 ± 8.19	9.32 ± 8.33	5.63 ± 2.29	5.29 ± 1.30

tracking. Note that the proposed approaches, CDP and TDP, exhibit better performance (visually and quantitatively) than direct interpolation, Generalized Wishart processes (GWP) and log-Euclidean interpolation. One would expect CDP and TDP to exhibit a similar performance and we can see that differences of Frobenius norm between CDP and TDP are not statistically significant.

Rank-4 and 6 results

Table 4.11 shows quantitative results of Frobenius distance for HOT interpolation in rank-4 and 6 synthetic data. Figure 4.22 shows the error map for rank-4. Figure 4.23 shows the histogram of generalized anisotropy (GA) for the rank-6 representation. As mentioned above, we only compare linear interpolation against the CDP and TDP due to lacking of methods for HOT interpolation.

TABLE 4.11: Frobenius distance for rank-4 and 6 synthetic HOT fields

	CDP	TDP	Direct interpolation
Rank-4	0.572 ± 0.468	0.464 ± 0.284	2.117 ± 2.551
Rank-6	1.386 ± 1.457	1.441 ± 1.614	2.845 ± 2.996

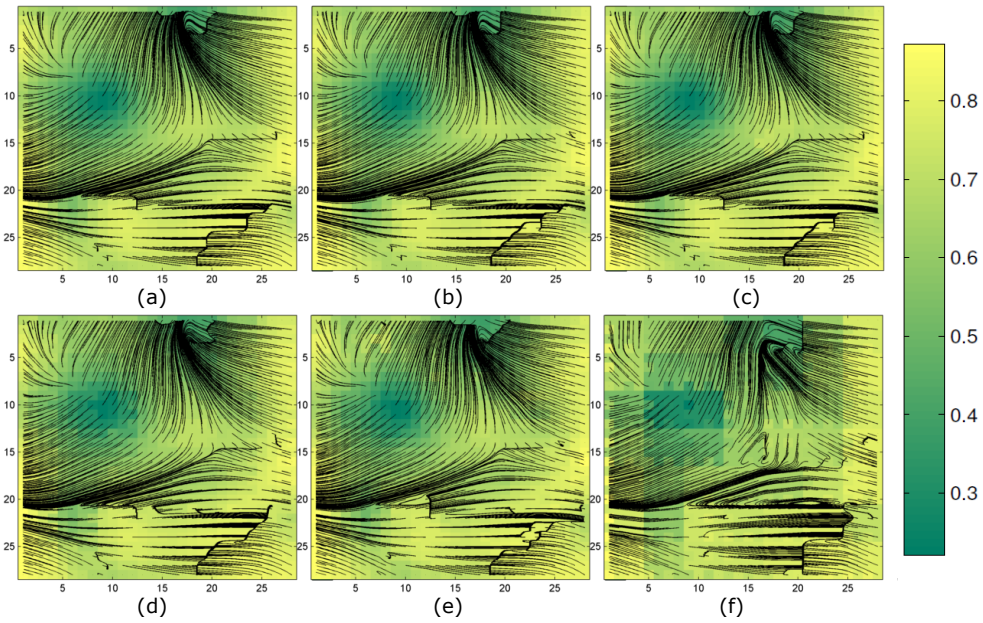


FIGURE 4.21: FA maps and fiber tracking of a rank-2 synthetic HOT field: (a) Ground-truth data, (b) CDP, (c) TDP, (d) Direct interpolation , (e) GWP, and (f) log-Euclidean.

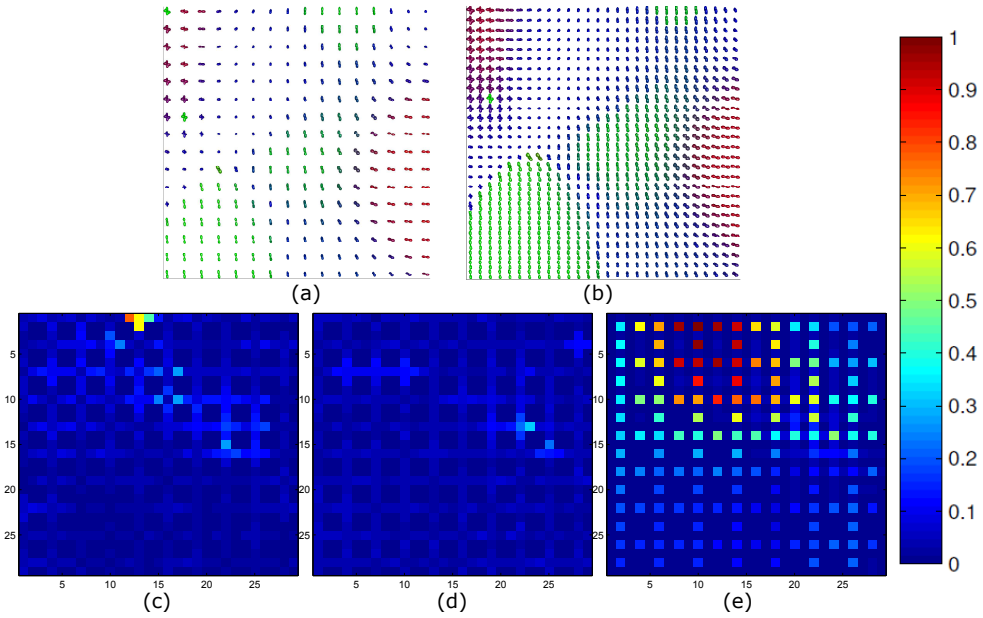


FIGURE 4.22: Normalized error maps for interpolation of rank-4 synthetic HOT field: (a) Training data, (b) Ground-truth, (c) CDP, (d) TDP, (e) linear interpolation.

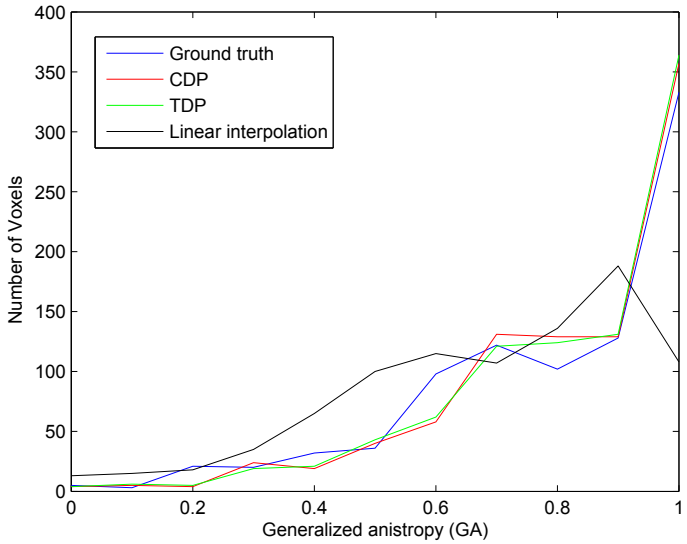


FIGURE 4.23: Generalized anisotropy (GA) histogram for the rank-6 synthetic HOT field

4.3.4 HOT fields interpolation in crossing fibers data

Rank-2 Results

Figure 4.24 and Table 4.12 show results for the rank-2 crossing fibers field. Figure 4.25 shows the FA maps and vectorial fiber tracking. Again, CDP and TDP outperform direct interpolation, Generalized Wishart processes (GWP) and log-Euclidean interpolation, when we evaluate accuracy and morphological validation.

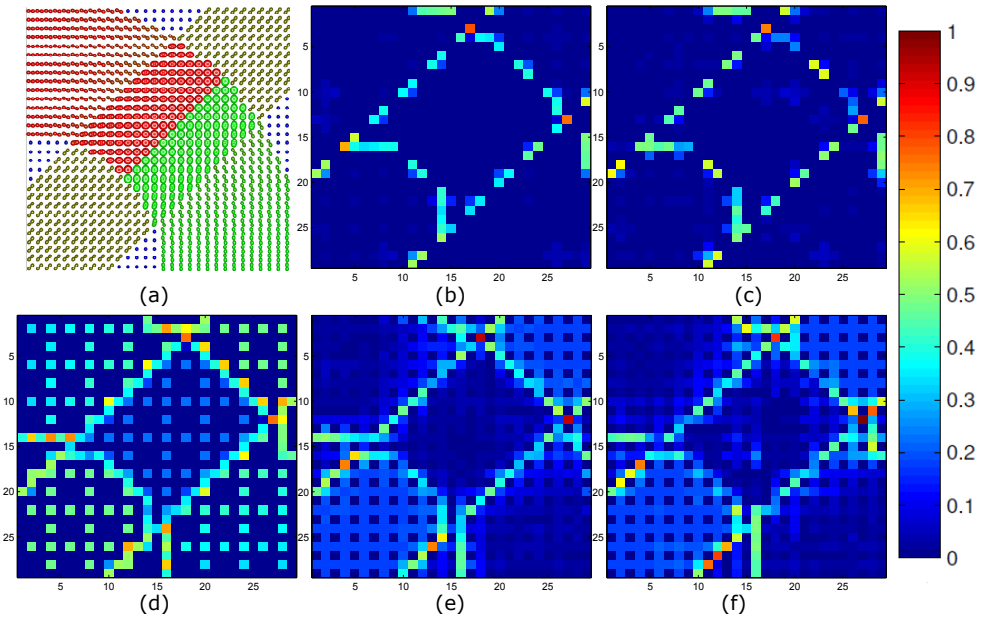


FIGURE 4.24: Normalized error maps for interpolation of a rank-2 crossing fibers HOT field: (a) Ground-truth, (b) CDP, (c) TDP, (d) Direct interpolation , (e) GWP, and (f) log-Euclidean.

TABLE 4.12: Frobenius distance and MSE of FA for rank-2 crossing fibers HOT field

	D. interpolation	Log-Euclidean	GWP	TDP	CDP
Frobenius distance.	0.551 ± 0.625	0.505 ± 0.531	0.372 ± 0.353	0.139 ± 0.112	0.125 ± 0.110
MSE of FA ($\times 10^{-3}$)	9.62 ± 3.34	9.53 ± 3.38	4.60 ± 2.30	3.90 ± 2.25	3.90 ± 2.36

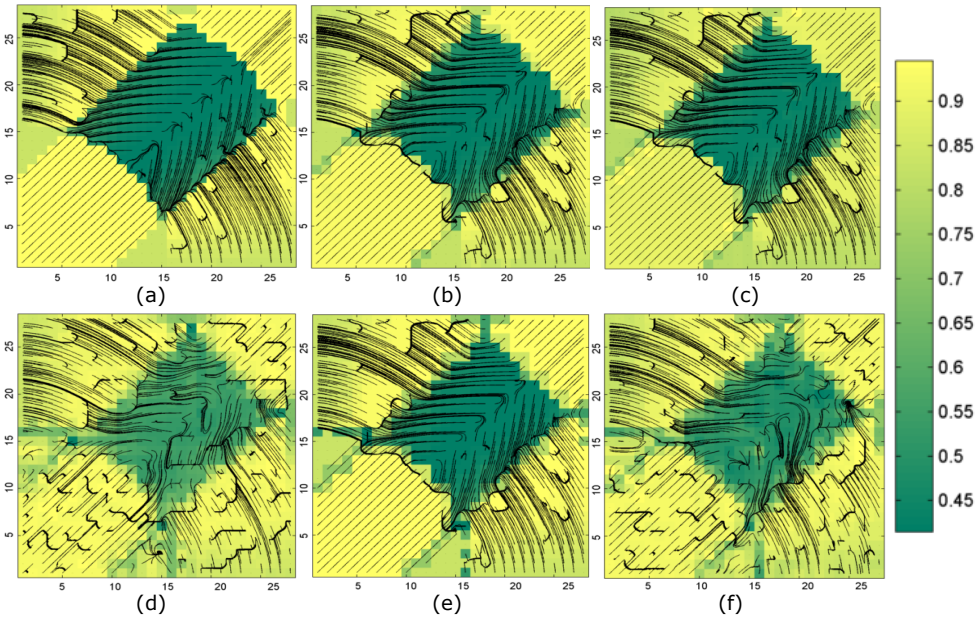


FIGURE 4.25: FA maps and fiber tracking of a rank-2 crossing fibers HOT field: (a) Ground-truth data, (b) CDP, (c) TDP, (d) Direct interpolation , (e) GWP, and (f) log-Euclidean.

Rank-4 and 6 results

TABLE 4.13: Frobenius distance for rank-4 and 6 crossing fibers HOT fields

	CDP	TDP	Direct interpolation
Rank-4	0.734 ± 0.634	0.790 ± 0.731	1.806 ± 1.957
Rank-6	1.214 ± 1.013	1.132 ± 0.994	2.976 ± 2.776

Table 4.13 show the average Frobenius distance for HOT interpolation in rank-4 and 6 crossing fibers data. Figures 4.26 and 4.27 show the error map for rank-6 and the histogram of generalized anisotropy (GA) for rank-4, respectively.

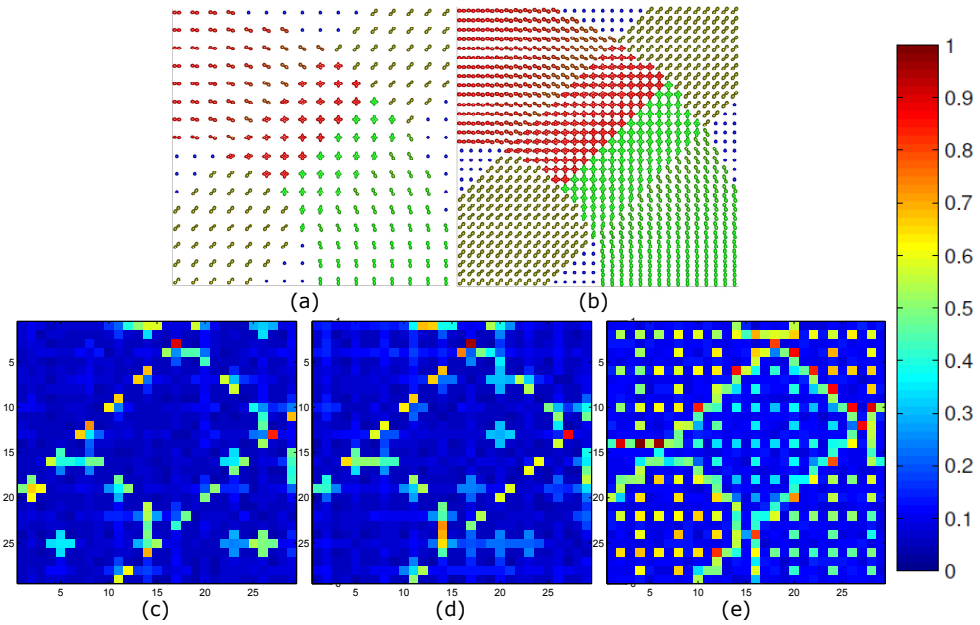


FIGURE 4.26: Normalized error maps for interpolation of rank-6 crossing fibers HOT field: (a) Training data, (b) Ground-truth, (c) CDP, (d) TDP, (e) linear interpolation.

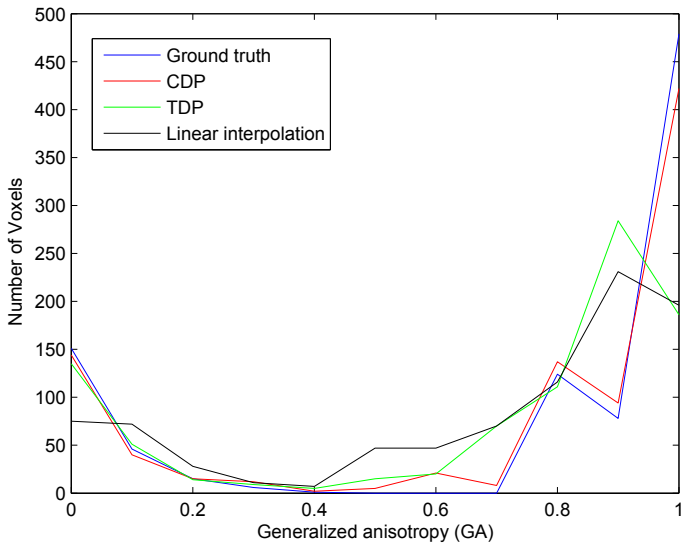


FIGURE 4.27: Generalized anisotropy (GA) histogram for the rank-4 crossing fibers HOT field

4.3.5 HOT fields interpolation in real dMRI data

Rank-2 Results

Figure 4.28 and Table 4.14 show results for the rank-2 real data. Figure 4.29 shows the FA maps and vectorial fiber tracking. Similar to the results for the synthetic and crossing fibers examples, the CDP and TDP offer better performance compared to the linear, log-Euclidean interpolation and GWP.

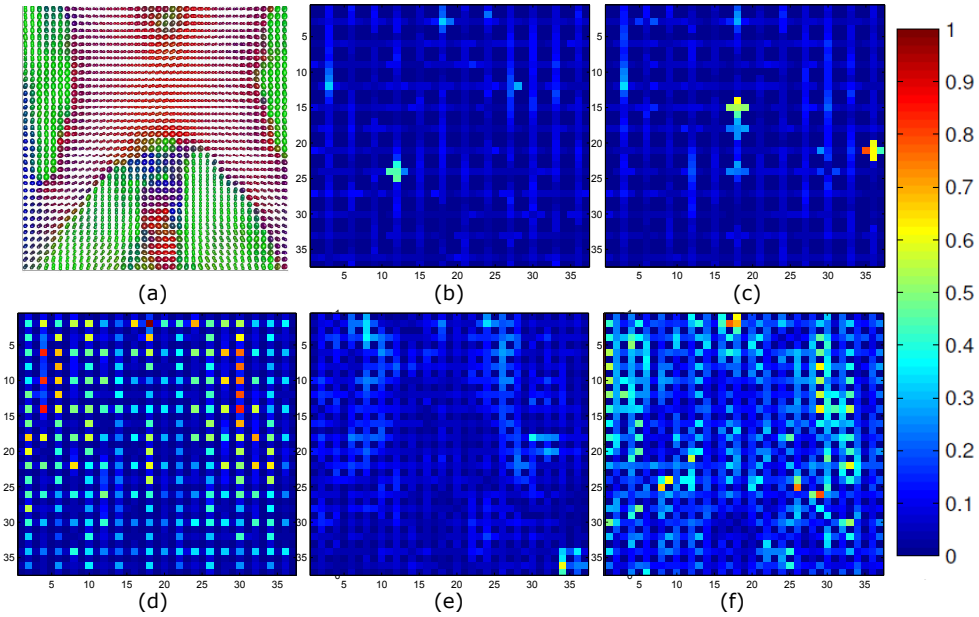


FIGURE 4.28: Normalized error maps for interpolation of rank-2 real HOT field: (a) Ground-truth data, (b) CDP, (c) TDP, (d) direct interpolation, (e) GWP, and (f) log-Euclidean.

TABLE 4.14: Frobenius distance for rank-2 real HOT field

	Direct interpolation	Log-Euclidean	GWP	TDP	CDP
Frobenius distance	0.275 ± 0.219	0.224 ± 0.196	0.182 ± 0.178	0.118 ± 0.105	0.102 ± 0.093
MSE of FA ($\times 10^{-3}$)	2.52 ± 1.91	2.47 ± 1.88	1.55 ± 0.81	1.42 ± 0.79	1.06 ± 0.35

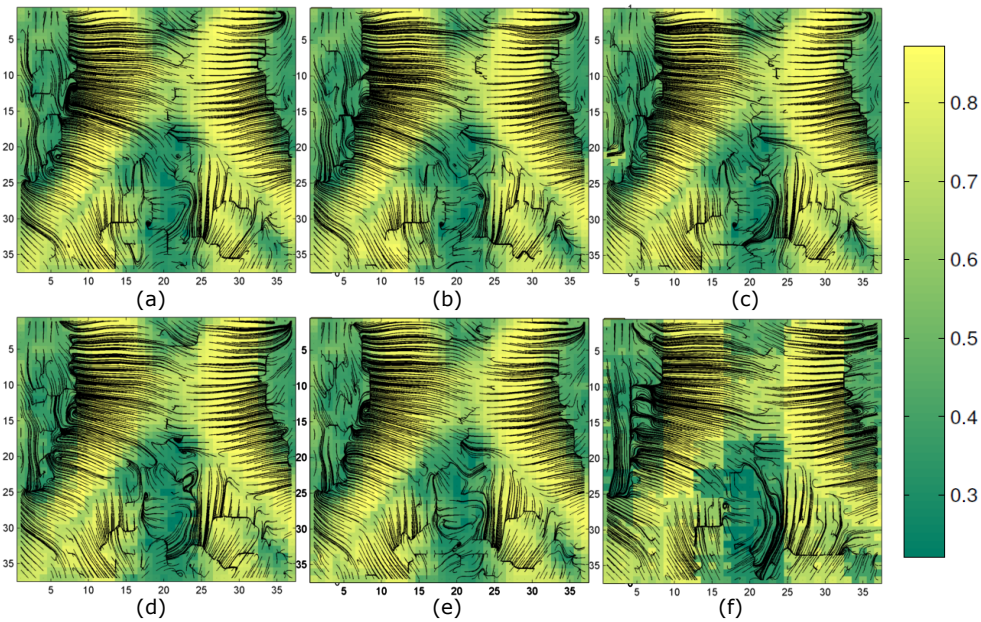


FIGURE 4.29: FA map and fiber tracking of a rank-2 real dMRI HOT field: (a) Ground-truth data, (b) CDP, (c) TDP, (d) Direct interpolation , (e) GWP, and (f) log-Euclidean.

Table 4.15 shows a quantitative comparison of the performance of methods in a 3D tractography carried out in the region of interest (ROI) centered in the corpus

callosum, having the number of generated fibers (NGF) and the average length of tracts (ALT) as metrics. Figure 4.30 shows the graphical comparison of the same procedure. Visual and quantitative comparison demonstrates that it is possible to improve the fibers tracts reconstruction through interpolation of tensor fields.

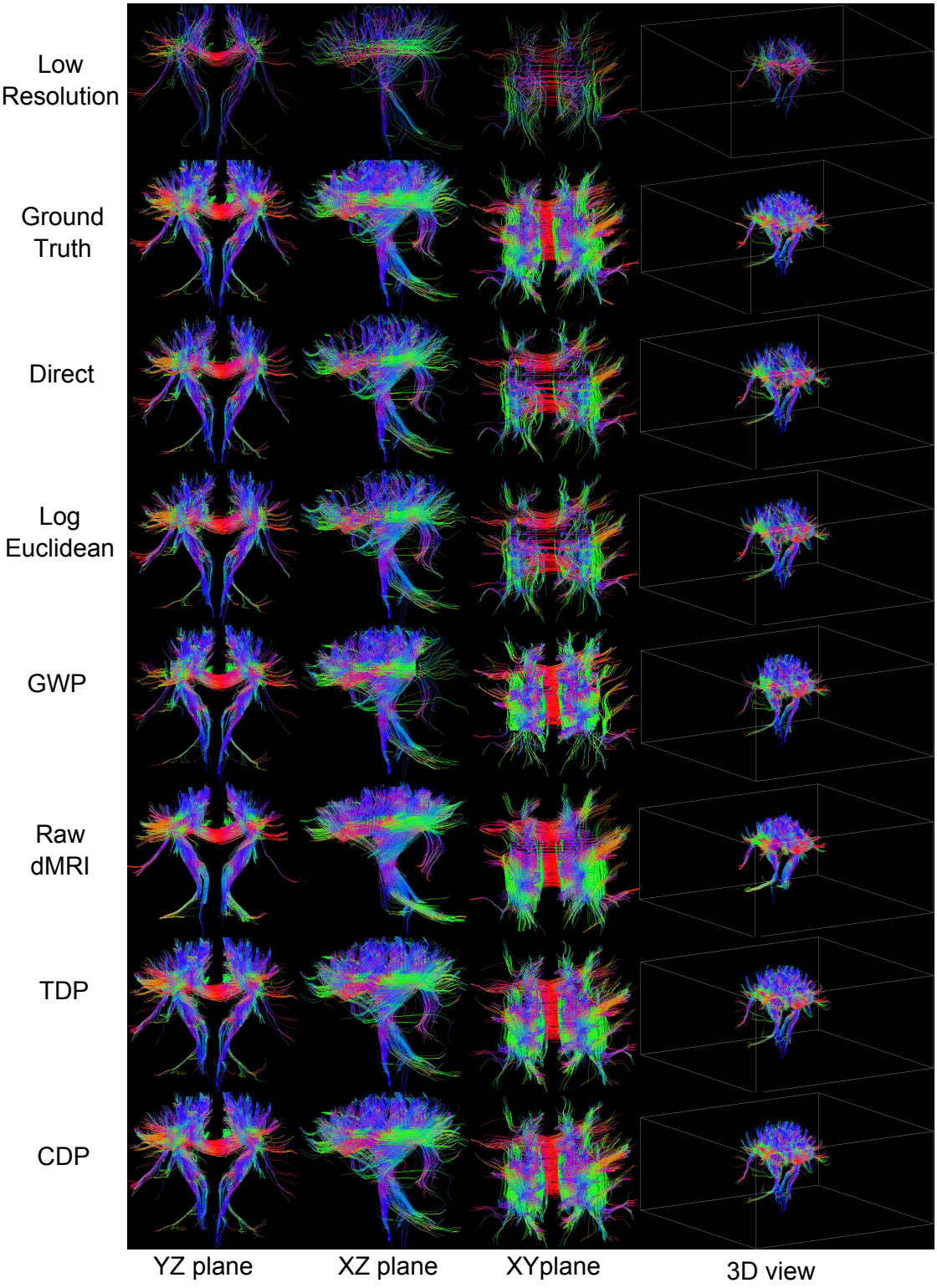


FIGURE 4.30: 3D tractography of the rank-2 dMRI field for the selected ROI centered in the corpus callosum. The reference is the ground-truth data. We show three cartesian planes and a 3D view.

TABLE 4.15: 3D Tractography metrics obtained by each interpolation method applied in the rank-2 real HOT field. NGF corresponds to the number of generated fibers and ALT is the average of length of tracts.

	NGF (Number)	Error NGF (%)	ALT(mm)	Error ALT (%)
Ground-Truth	2465	0	114.51	0
CDP	**2332**	**5.39**	113.66	0.74
TDP	2320	5.88	**113.85**	**0.58**
GWP	2236	9.29	112.6	1.67
Direct	1462	40.68	107.08	6.49
Log-Euclidean	1579	35.94	107.87	5.80
Raw-dMRI	1971	20.04	111.88	2.29

Rank-4 and 6 results

Table 4.16 show error results for HOT interpolation in rank-4 and 6 real dMRI data. Figures 4.31 and 4.32 illustrate the error map for rank-4 and the histogram of generalized anisotropy (GA) for rank-6, respectively. The TDP and CDP improve the performance when compared to linear interpolation and dMRI raw interpolation.

TABLE 4.16: Frobenius distance for rank-4 and 6 real HOT fields

	CDP	TDP	Raw dMRI	Direct interpolation
Rank-4	1.178 ± 1.025	1.320 ± 1.288	1.804 ± 0.978	2.739 ± 2.526
Rank-6	3.211 ± 2.923	3.492 ± 3.347	4.719 ± 2.547	6.243 ± 6.252

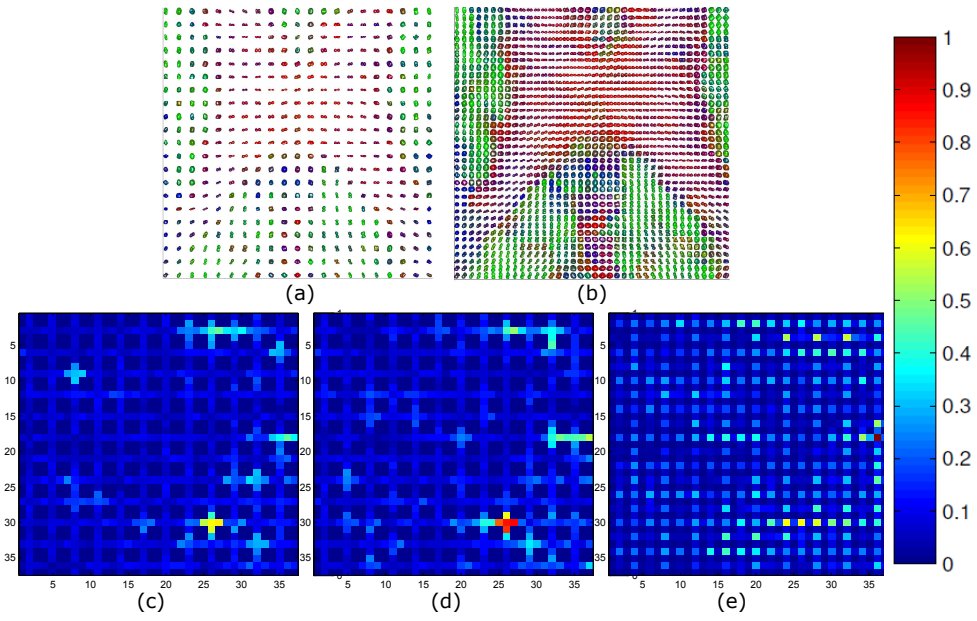


FIGURE 4.31: Normalized error maps for interpolation of rank-4 real HOT field: (a) Training data, (b) Ground-truth, (c) CDP, (d) TDP, (e) Raw dMRI

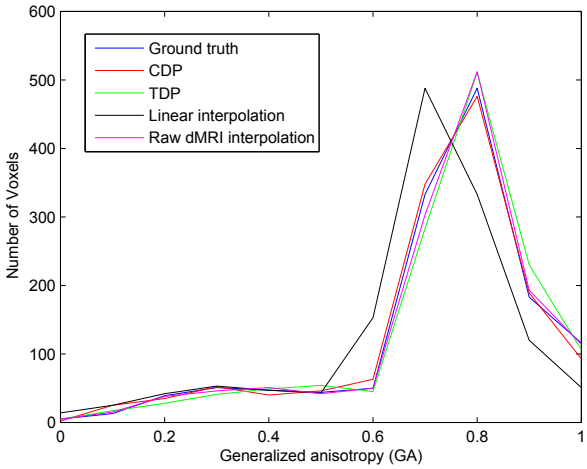


FIGURE 4.32: Generalized anisotropy (GA) histogram for the rank-6 real dMRI HOT field

4.3.6 Discussion

The proposed approaches (CDP and TDP) demonstrate better performance in interpolation of HOT fields of any order, compared to direct linear interpolation and dMRI raw interpolation. In rank-2 data, the CDP and TDP also outperform log-Euclidean interpolation and the recently proposed framework based on generalized Wishart processes. Our methods can capture the global spatial trend of the field. Thus, they can deliver a precise estimation of new data. The CDP and TDP are flexible to model several transitions inside HOT fields. This property is important because HOT data are very heterogeneous. Quantitative results of Frobenius distance of tables 4.10, 4.11, 4.12, 4.13, 4.14, 4.16 show that the CDP and TDP always outperform to the comparison methods, for each dataset and for any order. The accuracy in estimation of new data is mandatory for interpolation of HOT data. Another key factor is that the proposed methods ensure positive definite tensors.

The Rician noise analysis is very useful for probing robustness of the CDP and TDP. Diagrams of mean error and standard deviation of figures 4.15, 4.17, 4.19 show a better performance of CDP and TDP than the state of the art approaches. If we compare CDP and TDP each other, both methods obtain statistically similar results. This behavior remains constant for all evaluated cases of SNR levels (including an extreme case of SNR=1) and different tensor orders. Robustness to noise of CDP and TDP is due to probabilistic modeling (Gaussian processes) of their parameters. The GPs modulate those parameters considering the tensors as noisy data. Therefore, there is an assumption of intrinsic noise in the model. Unlike classical deterministic interpolation, a probabilistic inference methodology is not highly affected when the training data are corrupted by noise. We must consider that brain dMRI data are always altered by Rician noise and different artifacts added in acquisition procedure.

Qualitative results of figures 4.20, 4.22, 4.24, 4.26. 4.28, 4.31 illustrate an interesting behavior when there are strong changes among nearby tensors. Looking at the figures in detail, the traditional methods can not capture accurately the rapid transitions in the field, no matter the rank. The most extreme case is the crossing fibers fields (figures 4.24, 4.26). These strong changes in spatial dynamic of HOT fields are very difficult to follow, even for robust methods. The CDP and TDP capture with a low error these changes. For example the direct transition from blue to green tensors, when they are highly anisotropic. Probabilistic models presented

in this work adapt much better to abrupt changes compared to the other methods. Another remarkable aspect of the proposed approaches is the guarantee of positive definite (PD) tensors. For this reason, the estimated data are physically realizable. Some methods such as linear direct interpolation and log-Euclidean can not ensure the estimation of PD tensors in noisy data, i.e. the real dMRI dataset. FA maps illustrated in figures 4.21, 4.25, 4.29 and MSE of FA displayed in tables 4.10, 4.12, 4.14 demonstrate that probabilistic approaches for tensorial interpolation are robust and can preserve morphological properties relevant in clinical applications. Again, outcomes for CDP and TDP are better than the comparison methods in all experiments. A closer look in fiber projections (black lines in figures 4.21, 4.25, 4.29) show that tractography obtained with CDP and TDP is coherent, and there are not outliers or missing fibers.

3D tractography results exhibited in table 4.15 and figure 4.30 are particularly relevant to remark the pertinence of enhancing spatial resolution in HOT fields. Regarding this, interpolation of tensor data allows to highlight anatomical details that can be seen only in very high resolution acquisitions. We consider that a HOT study with enhanced spatial resolution can improve the quality of tractography and aids the mapping of tissue structures. If we observe, the fiber tracts reconstructed from low resolution data (see first row of figure 4.30) is poorer than the reconstruction from high resolution data (second row of figure 4.30), where the density and number of fibers is clearly inferior in the corpus callosum and surrounding regions. Additionally, interpolation of the HOT fields reveals more fined structural features of complex fiber bundles (i.e. crossing and bifurcated fibers), improves the representation of tract shapes, and it augments the contrast in tissue boundaries. Furthermore, segmentation of gray matter (GM) and white matter (WM) is easier from FA maps, because of the increased contrast. In this sense, enhancing spatial resolution of HOT data takes relevance in clinical applications. For example; the surgical planning, where tractography is employed to map the displacements of projecting tracts and low resolution scans can difficult the accuracy of this procedure [31]. Also, the preliminary diagnosis of tumorous, ischemic or inflammatory lessons of the spinal cord [79], where it is required high resolution visual data.

In agreement with quantitative results of table 4.15, the CDP and TDP achieve tractography metrics close to the ground-truth study (the selected ROI has a size of $40 \times 40 \times 10$ voxels, centered in the corpus callosum). In this case we evaluate the number of generated fibers (NGF) and the average length of tracts (ALT). The GWP and dMRI-raw interpolation obtain acceptable results, while the deterministic methods (direct and log-Euclidean) have the lower performance. Qualitative results of fiber reconstruction (see figure 4.30) show missing fibers and a considerable reduction of the fiber density in some regions for direct and log-Euclidean approaches. Moreover, we observe a smoothing of fiber tracts that generates a lose of contrast for dMRI-raw interpolation; the problem of this blurring effect is that tiny brain structures and edges tend to disappear. As we explained before, proposed methods interpolate the tensors with low error and obtain tractography metrics nearby to the gold-standard. Summarizing, when we employ probabilistic methodologies for interpolation of tensor data, it is possible to get accurate 3D tractographic reconstruction from post processed low resolution dMRI scans.

Generalized anisotropy is an extension of FA For rank-4 and 6. The GA histograms (figures 4.23, 4.27, 4.32) obtained for the proposed methods follow the trend of the ground truth, especially the CDP. It means that interpolation of HOT with CDP or TDP does not affect the intrinsic physiological information of dMRI. On the other hand, the linear interpolation can not retain the trends in high values of GA (from 0.8

to 1.0). Strong transitions (i.e crossing fibers) in HOT fields occur in anisotropic regions. In consequence, linear interpolation is not able to capture with good accuracy the complex tissue structures. The method of Raw dMRI interpolation tested only in the real dataset, can keep the GA tendency. Nonetheless, it generates swelling effect in the estimated tensors.

CDP and TDP are Bayesian models whose parameters are modeled with Gaussian processes. Due to the complexity of canonical and Tucker decomposition, it is not possible to achieve an analytical solution for the posterior distribution. It is necessary to employ Markov Chain Monte Carlo methods for finding the posterior (elliptical slice sampling and Metropolis Hastings). There is a bottleneck in the construction of tensors with Tucker or canonical decomposition, when we calculate Kronecker products. A feasible option is to employ *mode-n* products. However, this procedure demands more operations and loops. Therefore, computational cost of CDP and TDP increases considerably with the number of training data. As we pointed out before, we split the entire field in several subfields and we employ a patch-based scheme. Then, we execute the algorithms at the same time, improving the performance and reducing the time needed for a successful training.

An important aspect that must be taken into account is the slow convergence of the MCMC methods used in our framework. This issue is critical for a large number of training data, because the time demanded to complete the training stage increases considerably. For this reason, we split the entire field in subfields that we process simultaneously. Also, it is necessary to employ dynamic patches into each subfield (i.e. 3×3 for enhancing to 5×5).

When we make a direct comparison between CDP and TDP, we do not find statistically significant differences in their outcomes, no matter the rank or dataset. We can say both methods have a comparable performance under any condition. Intuitively, we think this identical performance of CDP and TDP is due to their similar mathematical construction based on outer products, taking account that parameters of CDP are scaled-vectors and parameters of TDP are matrices and a core tensor. Also, the MCMC-based learning stage is almost the same for them. In relation to the convergence of each proposed method, we observe that time of training stage of CDP is more affected by the tensor rank than TDP. As we explained before, the size of matrices of TDP does not depend of the rank, while the number of vectors and eigenvalues in CDP increases considerably for higher orders. For example, in rank-2 fields, the learning stage of CDP is faster than TDP. For rank-4 is similar. However, for rank-6 tensors there is a considerable difference in favor of TDP. According to this, we consider that CDP is a suitable approach for lower ranks (2 and 4), and TDP is the appropriated method for higher orders (6 or more).

Chapter 5

Conclusions and future works

5.1 Specific aim 1

To develop a stochastic methodology for super-resolution of anatomical MRI (T1 y T2) through scalar stochastic processes for supporting tissue segmentation.

- We presented a methodology for resolution enhancement of anatomical magnetic resonance images (T1 y T2) based on supervised learning. We modeled a Gaussian process regression for 2D multi-slice images. In order to enhance edge regions, we applied a 2D filtering stage. Our proposal outperforms to B-spline interpolation [20], nearest-neighbor [46], and the GPR method developed in [75]. Our methodology performs better in both high resolution (HR) reconstruction images and segmentation of white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF). These outcomes make our proposed method a promising approach to enhance spatial resolution in brain MR images.

  As future work, the methodology can be extended to perform 3D HR-MRI studies in order to estimate voxels corresponding to new slices. Also, we pretend to improve edge enhancement by using adaptive filters.

5.2 Specific aim 2

To build a stochastic interpolation approach applied in rank-2 tensor fields for enhancing spatial resolution in diffusion tensor images (DTI) and preserving fractional anisotropy of diffusion images. Application to fiber tracts reconstruction

- We developed a probabilistic methodology to interpolate Diffusion Tensor Imaging (DTI) data. We model a DTI field as a Generalized Wishart process (GWP). We employ approximate Bayesian inference for optimizing the relevant variables in GWP. Results obtained with GWP in synthetic and real DTI data outperform to commonly used geometric methods: linear [18] and log-Euclidean [19]. Also, our proposed method guarantees positive definite tensors and it avoids an issue in tensorial interpolation known as swelling effect.

  As future work, we would like to extend this concept describe tensorial data from crossing fibers.

- We presented a feature-based methodology for interpolation of diffusion tensor fields. We decompose the tensors in eigenvalues and Euler angles for performing multi-output regression with Gaussian processes (MOGP). Results obtained with MOGP on synthetic and real dMRI data outperforms to the classical log-euclidean method [19] and the feature-based scheme with linear regression [23]. We evaluated accuracy of interpolation (Frobenius norm and Riemann distance) and preservation of fractional anisotropy (FA). MOGP guarantees positive definite tensors, preserves FA of DTs and it avoids the swelling effect. Also, we observed that probabilistic approaches have a better performance than classical geometric methods.

  As future work, we propose a validation with electrical conductivity models for deep brain stimulation.

- We presented a probabilistic methodology for interpolation of diffusion tensor fields using a multi-output Gaussian processes with a non-stationary (NmoGp) kernel function. The proposed model decomposes the tensors in six features describing the main properties of a diffusion tensor (direction and orientation). We index the extracted features in spatial coordinates to interpolate them. The structure of the kernel function employed in this method combines several kernels with different properties. The purpose is to characterize complex fields such as crossing fibers. In this context, the non-stationary kernel allows differentiating between strong and uniform transitions. In this way, the interpolation of new tensors is more accurate in comparison to ordinary approaches. We tested NmoGp against state of the art methods: [19, 23, 38] in two different datasets: simulation of crossing fibers, and a dMRI segment. Outcomes proved that proposed model outperforms to the comparison methods evaluating accuracy with Frobenius and Riemann distances. Although, our proposed approach is accurate and robust for interpolation of complex tensor fields, there is an important issue: initialization of parameters and hyperparameters is not straightforward. Currently, we use a cross-validation procedure.

  As future work, we would like to extend non-stationary kernel functions to more complex models such as generalized Wishart processes.

5.3 Specific aim 3

To formulate a probabilistic approach for interpolation of higher order tensors (HOT) fields allowing the enhancement of spatial resolution for construction of generalized anisotropy maps.

- We introduced two novel methods for tensorial interpolation of diffusion magnetic resonance imaging: the canonical decomposition process (CDP) and the Tucker decomposition process (TDP). The proposed methods generalize to higher order tensors, in contrast to traditional methods presented in the state of the art, valid only for rank-2 tensors [18, 19, 55, 40]. The canonical and the Tucker process outperformed to linear method [18], log-Euclidean [19], Generalized Wishart processes [40], and dMRI raw interpolation [55], when we tested three different datasets and for tensor fields of rank-2, 4 and 6.

  Also, we performed a morphological validation. For rank-2 tensor fields we evaluated fractional anisotropy (FA) maps and tractography (2D and 3D). For rank-4 and 6 tensors, we obtained the generalized anisotropy (GA) histograms.

CDP and TDP can preserve morphological properties of dMRI, avoiding non positive definite tensors and the swelling effect. For HOT data, it was possible to achieve high accuracy in GA curves, even in anisotropic regions.

CDP and TDP are Bayesian models, where their parameters are defined by a set of Gaussian processes. The probabilistic nature of proposed approaches favored the robustness, flexibility, generalization capability, and adaptability to heterogeneous or noisy data. On the other hand, the comparison methods reduced considerably their performance in presence of high levels of Rician noise.

As future work, we are interested in alternative and faster techniques than MCMC to calculate the posterior distribution of CDP and TDP.

Bibliography

[1] D. Le Bihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanis, and M. Laval-Jeantet, "Mr imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders," *Radiology*, vol. 161, no. 2, pp. 401–407, 1986.

[2] D. Essen and et. al., "The human connectome project: A data acquisition perspective." *NeuroImage*, vol. 62, no. 4, pp. 2222–2231, 2012.

[3] A. Smilde, R. Bro, and P. Geladi, "Multi-way analysis: Applications in the chemical sciences," *Wiley, West Sussex*, 2004.

[4] M. Vasilescu and D. Terzopoulos, "Multilinear image analysis for facial recognition," *Proc. Int. Conf. Pattern Recognit (ICPR)*, vol. 2, pp. 511–514, 2002.

[5] M. Morup, L. Hansen, S. Arnfred, L. Lim, and K. Madsen, "Shift invariant multilinear decomposition of neuroimaging data," *NeuroImage*, vol. 42, no. 4, pp. 1439–1450, 2008.

[6] J. Isaac and J. Kulkarni, "Super resolution techniques for medical image processing," in *International Conference on Technologies for Sustainable Development*, 2015, pp. 1–6.

[7] E. Carmi, S. L., A. Noga, F. Amos, and D. Fia, "Resolution enhancement in MRI," *Magnetic Resonance Imaging*, vol. 24, pp. 133–154, 2006.

[8] W. Hollingworth, C. Todd, M. Bell, Q. Arafat, S. Girling, K. Karia, and D. A.K., "The diagnostic and therapeutic impact of MRI: an observational multi-centre study," *Clin Radiol*, vol. 55, no. 11, pp. 825–831, 2000.

[9] A. Hefnawy, "An efficient super-resolution approach for obtaining isotropic 3-D imaging using 2-D multi-slice MRI," *Egyptian Informatics Journal*, vol. 14, pp. 117–123, 2013.

[10] A. Patti, M. Sezan, and A. Teklap, "High-resolution image reconstruction from a low-resolution image sequence in the presence of time-varying motion blur," in *Proc. ICIP*, 1994, pp. 343–337.

[11] H. Greenspan, O. G., N. Kiryati, and S. Peled, "MRI inter-slice reconstruction using super resolution," *Magnetic Resonance Imaging*, vol. 20, pp. 437–446, 2002.

[12] A. Rueda, N. Malpica, and E. Romero, "Single-image super-resolution of brain MR images using overcomplete dictionaries," *Medical Image Analysis*, vol. 17, no. 1, pp. 113–132, 2012.

[13] K. Jafari-Khouzani, "MRI upsampling using feature-based nonlocal means approach," *IEEE Transactions on Medical Imaging*, vol. 30, no. 10, pp. 1969–1985, 2014.

[14] Y. Wang, J. Qiao, L. Jun-bao, A. Ping-Fu, C. Shu-Chuan, and J. Rodd, "Sparse representation-based MRI super-resolution reconstruction," *Measurement*, vol. 47, pp. 946–953, 2012.

[15] P. Basser, J. Mattiello, and D. Le Bihan, "Estimation of the effective self-diffusion tensor from the nmr spin echo," *J. Magn. Reson*, vol. 103, pp. 247–254, 1994.

[16] P. Basser, R. Mattiello, and D. Le Bihan, "Diffusion tensor echo-planar imaging of human brain," in *Proceedings of the SMRM*, vol. 56, 1993, pp. 584–562.

[17] I. Hotz, J. Sreevalsan-Nair, and B. Hamann, "Tensor field reconstruction based on eigenvector and eigenvalue interpolation," in *Scientific Visualization: Advanced Concepts*, 2010, pp. 110–123.

[18] S. Pajevic, A. Aldroubi, and P. Basser, "A continuous tensor field approximation of discrete dt-mri data for extracting microstructural and architectural features of tissue," *J. Magn. Reson.*, vol. 154, pp. 85–100, 2002.

[19] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Log-euclidean metrics for fast and simple calculus on diffusion tensors," *Magn. Reson. Med.*, vol. 56, pp. 411–421, 2006.

[20] A. Barmpoutis, B. Vemuri, T. Shepherd, and J. Forder, "Tensor splines for interpolation and approximation of dt-mri with applications to segmentation of isolated rat hippocampi," *IEEE TRANSACTIONS ON MEDICAL IMAGING*, vol. 26, pp. 1537–1546, 2007.

[21] P. Fletcher and S. Joshi, "Riemannian geometry for the statistical analysis of diffusion tensor data," *Signal Processing*, vol. 87, pp. 250–262, 2007.

[22] G. Kindlmann, R. Estepar, M. Niethammer, S. Haker, and C. Westin, "Geodesic-loxodromes for diffusion tensor interpolation and difference measurement," *Signal Processing*, vol. 10, pp. 1–9, 2007.

[23] F. Yang, Y. Zhu, I. Magnin, J. Luo, P. Croisille, and P. Kingsley, "Feature-based interpolation of diffusion tensor fields and application to human cardiac dt-mri," *Medical Image Analysis*, vol. 16, pp. 459–481, 2012.

[24] E. Ozarslan and T. Mareci, "Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging," *Magnetic Resonance in Medicine*, vol. 50, pp. 955–965, 2003.

[25] S. Mori, B. Crain, V. Chacko, and P. van Zijl, "Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging," *Ann Neurol*, vol. 45, pp. 265–269, 1999.

[26] C. Liu, R. Bammer, B. Acar, and M. Moseley, "Characterizing non-gaussian diffusion by using generalized diffusion tensors," *Magnetic Resonance in Medicine*, vol. 51, pp. 924–937, 2004.

[27] M. Moakher, "Fourth-order cartesian tensors: old and new facts, notions and applications," *The Quarterly Journal of Mechanics and Applied Mathematics*, vol. 61, pp. 181–203, 2008.

[28] A. Barmpoutis and B. Vemuri, "A unified framework for estimating diffusion tensors of any order with symmetric positive-definite constraints," in *Proceedings of ISBI10: IEEE International Symposium on Biomedical Imaging*, 2010, pp. 1385–1388.

[29] L. Astola and L. Florack, "Finsler geometry on higher order tensor fields and applications to high angular resolution diffusion imaging," *Lecture Notes on Computer Science*, vol. 5567, pp. 224–234, 2009.

[30] L. Astola, A. Jalba, E. Balmashnova, and L. Florack, "Finsler streamline tracking with single tensor orientation distribution function for high angular resolution diffusion imaging," *J Math Imaging Vis.*, vol. 41, pp. 170–181, 2011.

[31] T. Dirby, H. Lundell, M. Burke, N. Reislev, O. Paulsona, M. Ptito, and H. Siebner, "Interpolation of diffusion weighted imaging datasets," *Neuroimage*, vol. 103, pp. 202–213, 2014.

[32] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA: The MIT Press, 2006.

[33] K. Malczewski, "Inter-k-space motion based strategy for super-resolution in mri," in *17th European Signal Processing Conference*, 2009, pp. 30–34.

[34] I. Yassine and T. McGraw, "4th order diffusion tensor interpolation with divergence and curl constrained bezier patches," in *Proceedings of ISBI09: IEEE International Symposium on Biomedical Imaging*, 2009, pp. 634–637.

[35] P. Kroonenberg, *Applied multiway data analysis*. Wiley, 2008.

[36] N. Sidiropoulos, R. Bro, and G. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Signal Process.*, vol. 48, pp. 2377–2388, 2004.

[37] H. Vargas Cardona, A. Lopera, M. Alvarez, A. Orozco, M. N., and J. Tamames, "Gaussian processes for slice-based super-resolution MR images," *Lecture Notes in Computer Science*, vol. 9475, pp. 692–701, 2015.

[38] H. D. V. Cardona, Á. A. Orozco, and M. A. Álvarez, "Multi-output gaussian processes for enhancing resolution of diffusion tensor fields," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*. IEEE, 2016, pp. 1111–1114.

[39] J. Cuellar, M. Vargas Cardona, H.D. Alvarez, A. Álvarez, and A. Orozco, "Non-stationary multi-output gaussian processes for enhancing resolution over diffusion tensor fields," *Lecture Notes in Computer Science*, vol. 10657, pp. 168–176, 2017.

[40] H. Vargas Cardona, M. Alvarez, and A. Orozco, "Generalized Wishart processes for interpolation over diffusion tensor fields," *Lecture Notes in Computer Science*, vol. 9475, pp. 499–508, 2015.

[41] H. Vargas-Cardona, M. Alvarez, and A. Orozco, "Estimación de la propagación eléctrica cerebral generada por la dbs en pacientes con enfermedad de parkinson de la región sur-occidente de colombia," *Ingeniería y Desarrollo*, vol. 34, pp. 117–138, 2016.

[42] H. Vargas Cardona, M. Alvarez, and A. Orozco, "Analysis of the geometry and electric properties of brain tissue in simulation models for deep brain stimulation," *Lecture Notes in Computer Science*, vol. 10125, pp. 493–501, 2017.

[43] T. Huang and T. R.Y., "Multi-frame image restoration and registration," *Advances in computer vision and image processing*, vol. 1, no. 7, pp. 317–339, 1984.

[44] S. Kim, N. Bose, and H. Valenzuela, "Recursive reconstruction of high resolution image from noisy undersampled frames," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, pp. 1013–1027, 1990.

[45] M. Aguena, N. Mascarenhas, J. Anacleto, and S. Fels, "MRI iterative super resolution with Wiener filter regularization," in *XXVI Conference on Graphics, Patterns and Images*, 2013, pp. 155–162.

[46] Y. Lu and R. Yang, "Super-resolution reconstruction of dynamic MRI by patch learning," in *12th International Conference on Control, Automation, Robotics and Vision*, 2012, pp. 1443–1448.

[47] F. Rousseau and C. Studholme, "A supervised patch-based image reconstruction technique: Application to brain MRI super-resolution," in *10th International Symposium on Biomedical Imaging*, 2013, pp. 346–349.

[48] A. Jog, A. Carass, and J. Prince, "Improving magnetic resonance resolution with supervised learning," in *IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, 2014, pp. 987–990.

[49] D. Le Bihan, F. Mangin, C. Poupon, C. Clark, S. Pappata, N. Molko, and H. Chabriat, "Diffusion tensor imaging: concepts and applications," *J. Magn. Reson. Imag.*, vol. 13, pp. 534–546, 2001.

[50] P. Basser, "Inferring microstructural features and the physiological state of tissues from diffusion weighted images," *NMR Biomed*, vol. 8, pp. 333–344, 1995.

[51] D. Tuch, V. Weeden, and A. Dale, "Conductivity tensor mapping of the human brain using diffusion tensor MRI," *Proc Natl Acad Sci*, vol. 98, pp. 11 697–11 701, 2001.

[52] H. Lombaert, J.-M. Peyrat, P. Croisille, S. Rapacchi, L. Fanton, P. Clarysse, H. Delingette, and N. Ayache, "Statistical analysis of the human cardiac fiber architecture from DT-MRI," in *Conf Proc FIMH*, 2011, pp. 171–179.

[53] C. Frindel, M. Robini, P. Croisille, and Y.-M. Zhu, "A graph-based approach for automatic cardiac tractography," *Magn. Reson. Med.*, vol. 64, no. 4, pp. 1215–1229, 2010.

[54] C. Bi, S. Takahashi, and I. Fujishiro, "Interpolating 3d diffusion tensors in 2d planar domain by locating degenerate lines," in *ISVC'10 Proceedings of the 6th international conference on Advances in visual computing*, M. Abadi and T. Ito, Eds., vol. 1. Springer-Verlag, 2010, pp. 328–337.

[55] F. Yang, Y. Zhu, J. Luo, M. Robini, J. Liu, and P. Croisille, "A comparative study of different level interpolations for improving spatial resolution in diffusion tensor imaging," *IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS*, vol. 18, no. 4, pp. 1317–1327, 2014.

[56] I. Yassine and T. McGraw, "A subdivision approach to tensor field interpolation," in *Workshop On Computational Diffusion MRI*, 2008, pp. 117–124.

[57] J. Jin, *Electromagnetic Analysis and Design in Magnetic Resonance Imaging*, 1st ed. CRC Press, 1997.

[58] J. Jin, J. Chen, and Z. Feng, "Numerical simulation of sar and b1-field inhomogeneity of shielded rf coils loaded with the human head," *IEEE Transactions On Biomedical Engineering*, vol. 45, no. 5, pp. 650–659, 1998.

[59] S. Jen, J. Chen, F. Lin, and W. Kuan, "Quantitative spectral/spatial analysis of phased array coil in magnetic resonance imaging based on method of moment," *IEEE Transactions On Medical Imaging*, vol. 18, no. 12, pp. 1129–1137, 1999.

[60] I. Tamer, Y. Lee, K. Allahyar, L. Pierre-Marie, A. Baertlein, and Z. Zlem, "Theoretical model for an mri radio frequency resonator," *IEEE Transactions On Biomedical Engineering*, vol. 47, no. 4, pp. 535–546, 2000.

[61] A. Webb and N. B. Smith, *Introduction to Medical Imaging: Physics, Engineering and Clinical Applications*, 1st ed. CAMBRIDGE University Press, 2011.

[62] G. Vecchi, E. Attardo, and T. Isernia, "Field synthesis in inhomogeneous media: Joint control of polarization, uniformity and sar in mri b1-field," *Progress In Electromagnetics Research*, vol. 118, pp. 355–377, 2011.

[63] B. Baertlein, A. Abduljalil, H. Zhu, R. P.M., S. Tamer, and L. I.R., "Effect of rf coil excitation on field inhomogeneity at ultra high

fieelds: A

field optimized tem resonator," *Magnetic Resonance Imaging*, vol. 19, pp. 1339–1347, 2001.

[64] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.

[65] A. Wilson and Z. Ghahramani, "Generalised wishart processes," in *UAI-11*, 2011, pp. 736–744.

[66] I. Murray, R. Adams, and D. Mackay, "Elliptical slice sampling," *JMLR*, vol. 9, pp. 541–548, 2010.

[67] M. Alvarez, L. Rosasco, and N. Lawrence, "Kernels for vector-valued functions: A review," *Foundations and Trends in Machine Learning*, vol. 4, pp. 195–266, 2011.

[68] W. Herlands, A. Wilson, H. Nickisch, S. Flaxman, D. Neill, W. Van Panhuis, and E. Xing, "Scalable gaussian processes for characterizing multidimensional change surfaces," *arXiv preprint arXiv:1511.04408*, 2015.

[69] J. Carroll and J. Chang, "Analysis of individual differences in multidimensional scaling via n-way generalization of eckart-young decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, 1970.

[70] H. Gulliksen and N. Frederiksen, Eds., *Contributions to mathematical psychology*. New York: Holt, Rinehart and Winston, 1964.

[71] L. Qi, F. Wang, and Y. Wang, "Z-eigenvalue methods for a global polynomial optimization problem," *Mathematical Programming*, vol. 118, no. 2, pp. 301–316, 2009.

[72] P. Comon, G. Golub, L.-H. Lim, and B. Mourrain, "Symmetric tensors and symmetric tensor rank," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1254–1279, 2008.

[73] R. Hasrhman, "Foundations of the parafac procedure: Models and conditions for an explanatory multimodal factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, pp. 1–84, 1970.

[74] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods (Springer Texts in Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2005.

[75] H. He and W.-C. Siu, "Single image super-resolution using Gaussian process regression," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2011, pp. 449–456.

[76] R. Kikinis, S. Pieper, and K. Vosburgh, "3d slicer: A platform for subject-specific image analysis, visualization, and clinical support," in *Jolesz FA, editor. New York: Springer*, 2014, pp. 277–289. [Online]. Available: http://www.slicer.org/

[77] M. Alvarez and N. Lawrence, "Computationally efficient convolved multiple output gaussian processes," *Journal of Machine Learning Research*, vol. 12, pp. 1459–1500, 2011.

[78] C. Lin-Chin, D. Jones, and C. Pierpaoli, "RESTORE: Robust estimation of tensors by outlier rejection," *Magnetic Resonance in Medicine*, vol. 53, pp. 1088–1095, 2005.

[79] M. Vargas, J. Delavelle, H. Jlassi, B. Rilliet, M. Viallon, C. Becker, and L. K., "Clinical applications of diffusion tensor tractography of the spinal cord," *Neuroradiology*, vol. 50, pp. 25–29, 2008.